BANK OF ENGLAND

Staff Working Paper No. 983

Measuring Capital at Risk in the UK banking sector: a microstructural network approach

Giovanni Covi, James Brookes and Charumathi Raja

May 2022

BANK OF ENGLAND

Staff Working Paper No. 983

Measuring Capital at Risk in the UK banking sector: a microstructural network approach

Giovanni Covi,[1] James Brookes[2] and Charumathi Raja[3]

Abstract

In this paper we construct and analyse the UK banking system's Global Network of granular exposures which captures roughly 90% of the UK banking system's total assets for the period 2018 Q1 to 2021 Q4. We thus study the microstructure of UK banking system focusing on the role played by concentration risk and interconnectedness across sectors. We then estimate the quarterly evolution of expected losses (Capital at Risk) for the UK banking sector, and via Monte Carlo simulations the stochastic distribution of UK banks' losses to study the severity and likelihood of tail-events (Conditional Capital at Risk). In the end, we provide insights on the impact of the Covid-19 pandemic on UK banking system's loss distribution by decomposing the sources of average and tail risks.

Key words: Financial network, systemic risk, stress testing, Covid-19 pandemic.

JEL classification: D85, G21, G32, L14.

(1) Stress Test Strategy Division, Bank of England. Email: giovanni.covi@bankofengland.co.uk
(2) Advanced Analytics Division, Bank of England: Email: james.brookes@bankofengland.co.uk
(3) Banking Capital Policy Division, Bank of England: Email: charu.raja@bankofengland.co.uk

ISSN 1749-9135 (on-line)

"Mapping the networks between financial institutions is a first step towards gaining a better understanding of modern financial systems. A network perspective would not only account for the various connections within the financial sector or between the financial sector and other sectors, but would also consider the quality of these links. We need this work to guide the development of new theories that can help us understand events such as the August 2007 crisis, as well as design new regulations that better meet the challenge of an increasingly networked world" (Allen and Babus, 2009: 13).

1. Introduction

The Great Financial Crisis and its aftermath have highlighted how the functioning of the economic system is strongly intertwined with the functioning of the financial system. Input-output linkages among non-financial corporations affect the productive structure of our economic system, whereas contractual relationships among financial institutions determine the microstructure of the financial network. All in all, by providing funding through loan exposures and by allocating capital in the form of debt and equity security exposures to the real economy, the financial system shapes the financial-economic nexus, binding together the two systems' return and risk cycle. Hence, the way these contractual relationships are distributed within and between the two systems - interconnectedness and concentration risk - has relevant implications for business cycle fluctuations and financial stability. Following Allen and Babus (2009)'s call for action, by adopting a microstructural network perspective our work aims at empirically investigating one side of this nexus: how risk propagates from within the economic system to banks' balance sheet.

In this respect, Gabaix (2011) has showed how concentration risk in the economic system such as shocks to large non-financial corporations may lead to remarkable fluctuations in economic activity - the granular origins of aggregate fluctuations. Moreover, Acemoglu et al. (2012) showed how interconnectedness in economic activity such as a high level of interdependency in the intersectoral input-output firms' linkages - network origins - may explain aggregate fluctuations in output. These results are extremely relevant in light of the Covid-19 pandemic, which has caused bottleneck problems in the global supply chain (Rees and Rungcharoenkitkul, 2021). These network features - concentration risk and interconnectedness - also play an important role within the financial system determining fluctuations in the level of systemic risk. The relationship between the functioning of the financial system and that of the economic system and the implications for the financial system's stability have been explored by Acemoglu et al. (2015). According to this work, financial stability and the level of systemic risk in financial networks depend on the size of the initial shock stemming from the real economy and on the financial network structure. More dense (sparse) financial networks tend to amplify (dampen) real economic shocks when the initial shock is sizeable, whereas more dense financial networks enhance financial stability by reducing financial contagion when shocks are less sizeable.

The financial network literature (Glasserman and Young, 2016) has focused on modelling and quantifying those contagion channels at play that determine financial amplification mechanisms usually relying on an exogenous trigger like a bank default event. However, there has not been much emphasis in the literature on measuring the severity and probability of the trigger event. According to Acemoglu et al. (2015), it is this variable together with the network structure that determines the system's propensity to instability and thus the size of financial amplification mechanisms.

Leveraging-up these insights from the literature, which is reviewed in Section 2, our work empirically measures the level of systemic risk in the UK banking sector by applying a stochastic microstructural balance sheet-based network methodology on the UK banks' global network of granular exposures. We thereby quantifying the potential severity of the initial shock in monetary terms (£ amounts) – expected and tail losses - as well as the probability of systemic events. We contribute to the stress testing literature by performing a stochastic stress test for the UK banking sector in which we compute the UK banking system's loss distribution as a function of: i) the actual network structure of UK banks' exposures, thereby modelling interconnectedness and concentration risks; ii) a correlation matrix of counterparty defaults approximating Acemoglu et al. (2015)'s intersectoral input-output firms' linkages; and iii) a set of risk factors - loss given default and probability of defaults parameters – which captures UK banks' 1-year ahead view by country and sector of the state of the world economy. The novelty of this work is rooted in its stochastic approach to stress testing, thereby modelling scenario uncertainty, and in its empirical dimension, the estimation of the UK banks' loss distribution using a quasi-complete network perspective, and the related rich set of findings on the microstructure of UK banking system, on the impact of Covid-19 pandemic on UK banking system's tail risk, and on the decomposition of the sources of tail-risk.

To achieve this, we so construct the most comprehensive granular exposure-based dataset for the UK banking sector to date by merging different supervisory datasets available at the Bank of England. In this respect, the resulting Global Network of UK banks' loan, security and derivative exposures covers roughly £9.4 trillion or 90% of the UK banking system' asset side. Out of this £9.4 trillion of exposures, 43% is captured with a bank-to-counterparty relationship (granular exposures), whereas the remaining 57% is composed by aggregated exposure with a bank-to-country/sector relationship. Overall, the Global Network, which is updated quarterly, spans 16 quarters between Q1-2018 and Q4-2021, and comprises 36 UK reporting banks, 7200 counterparties and 17.500 contractual relationships distributed across six major sectors (non-financial corporates, financial corporates, credit institutions, governments, central banks, and households) and across more than 170 countries. Given that the technical details of the dataset's construction are quite involved, we provide a description of the dataset's construction at the beginning of Section 3 and a more detailed discussion in Appendix A for the interested reader.

With this newly constructed dataset at our disposal, the bulk of Section 3 and Section 4 contribute to the study on the microstructure of financial systems and on comparative financial systems led by the seminal paper of Allen and Gale (2001). We thus present detailed network statistics across time, sectors and exposure types in order to provide a footprint for the microstructure of the UK banking system and so moving a step forward in the mapping of the Global Banking System. Our key results are that the UK banking network is composed of: i) a core-double periphery structure, that is, an additional periphery which is an exclusive market for certain key players; ii) a highly fat-tailed distribution of exposures with roughly the top-10% of exposures and counterparties capturing 90% of the total exposure amounts, iii) and a quite interconnected banking system across sectors and countries.

Next, Section 5 exploits the microstructure of the Global Network to quantify UK banking system's *Capital at Risk* (CaR) and *Conditional Capital at Risk* (CCaR), that is, respectively potential expected losses and tail loss estimates before and during the Covid-19 pandemic. In this respect, the UK banking system's CaR estimate in Q2-2020 (peak of the Covid-19 crisis) was close to £37 billion, up by 36% (£10 billion) relative to pre-crisis average expected loss estimates.

Next we compute *Conditional Capital at Risk* (CCaR), that is, tail-loss estimates, so as to model scenario uncertainty and the severity of stress events as a function of concentration risk and interconnectedness of the UK banking sector's network of exposures. Overall, this approach allows us to assess not only the severity of extreme events, but also to measure their likelihood. In this respect, we estimate that an *extreme stress* scenario (99^{th} percentile of the estimated loss distribution) would produce roughly £147 billion of losses in Q2-2020 up from an average of £91 billion in the pre-pandemic period. Moreover, we estimate that due to the Covid-19 pandemic the likelihood of the UK banking sector experiencing an extreme stress event above £91 billion losses increases from 1% pre-crisis to 4.1% at the peak of the pandemic in Q2-2020 (one extreme stress every 24 years). By comparing the increase in the severity and likelihood of tail events with the increase in expected losses (CaR estimates) at the peak of the crisis (Q2-2020), we find that average risk increase (36%) by less than the increase in tail risk, respectively 310% and 62% for the probability and severity of extreme stress events. This finding highlights that expected losses (Capital at Risk) is not a good proxy for measuring tail risk (Conditional Capital at Risk).

In the end, we show that a tightening in the correlation structure of banks' counterparty defaults leads to an increased severity and likelihood of tail events, thereby corroborating empirically Acemoglu et al. (2015)'s findings. According to this stressed correlation structure, we estimate that the probability of experiencing an extreme stress event would increase to 6% (4.1%). In the end, we provide insights on the sources of average and tail risks by deriving the loss contribution of changes in PD and LGD parameters and the network structure. We show that the network structure may have a positive contribution to average risk, but at the same time the same network structure may have a negative contribution to tail risk.

Our paper concludes with Section 6, which provides a synopsis and contextualizes the contributions of the paper in relation to the existing literature.

2. Literature Review

Granular data collection such as the large exposures framework were introduced by the Basel Committee on Banking Supervision in order to measure and monitor concentration risk and interconnectedness of credit institutions (BCBS, 2014). The collection of these data fostered empirical research analysing financial networks, which before that, relied on synthetic bilateral linkages imputed using optimization methods such as maximum, minimum or relative entropy solutions based on firms' balance sheet information (Sheldon and Maurer 1998, Degryse and Nyguyen, 2007; Elsinger et al. 2006; Upper, 2011, Van Lelyveld and Liedorp, 2006), or by generating random networks consistent with partial information (Halaj and Kok, 2013; Anand et al., 2014).

As emphasized by Glasserman and Young (2016) empirical work in this field was and is still limited by the confidentiality of interbank transactions, which are available only to central bankers and supervisors. Given these data limitations, most of empirical analyses tend to be country-specific or market-specific and focused on studying interbank network relationships[1].For the UK financial network, our jurisdiction of interest, Gai et al. (2011) exploited the large exposure data to study concentration and contagion in the UK interbank market focusing on liquidity risk. Bardoscia et al. (2019a) exploited large exposure data and interbank security and derivative exposures to study how solvency contagion may propagate in the UK interbank market. Other UK-centric studies have investigated specific market

[1] See Huser (2015) and Bardoscia et al (2021) for a review of the financial network literature.

segments, such as Coen et al. (2019) who focused on UK banks' security holdings and fire-sales spillovers, and Huser et al. (2021) on interbank and CCPs repo market relationships in times of stress, as well as Bardoscia et al. (2019b) who studied the UK OTC derivative market.

In this respect, the construction of the Global Network dataset - to our knowledge the most comprehensive exposure-based dataset for the UK banking sector to date - allows us to complement this stream of literature: i) by adopting a multi-sector perspective, thereby going beyond contractual relationships in the interbank market; ii) by covering an extended sample of UK reporting banks, not limited to just the largest institutions as it was the case in previous studies; and iii) by embracing a multilayer perspective of loan, security and derivative exposures, not limited to a single market segment or a set of largest exposures, thus almost covering the complete asset side of the UK banking sector.

Thanks to our dataset, we contribute to the study on the microstructure of financial systems and on comparative financial systems. In this regard, we document the degree of interconnectedness between UK banks and financial corporates, non-financial sector corporations, governments and central banks as well as the degree of concentration risk on an exposure and counterparty level. We then provide a comparison between the UK and Euro area (EA) banking networks by leveraging the work of Montagna et al. (2021) and Sydow et al. (2021) on the EA banking sector, the only two studies that like ours have a similar granular multilayer coverage of banks' loan and security exposures across sectors and countries.

Next, we contribute to the stress testing literature by implementing a microstructural balance sheet-based network methodology with a stochastic component, which allows us to model scenario uncertainty, similarly to Montagna et al. (2021) and Sydow et al. (2021), and capture the role played by concentration risk and interconnectedness in the determination of tail-events, two features that are not modelled in a classical stress testing approach. Moreover, thanks to this stochastic approach to stress testing, we are able to quantify not only the severity of tail events, but also to estimate their likelihood. In contrast to the financial network literature (Glasserman and Young, 2016), the modelling of contagion and amplification mechanisms taking place in the interbank market is left aside as an extension for future work. This modelling choice allows us to focus the analysis on the role played by the distribution of exposures across sectors, instead of within the interbank network.

Next, we contribute to the systemic risk literature by deriving as output measures unconditional and conditional estimates of Capital at Risk, similar to the SRISK measure of Bronwlees and Engle (2017) as well as the Covar approach of Adrian and Brunnermeier (2019). In this respect, we follow Acemoglu et al. (2015)'s theoretical framework and measure the level of systemic risk in the UK banking sector in relation to the size of the initial shock stemming from exposures towards the real economy and the financial network structure. Thus, we test our results to a variation in the correlation structure of UK banks' counterparty defaults, thereby modelling explicitly intersectoral input-output firms' linkages as in Acemoglu et al. (2012). In this respect, the present work provides empirical evidence and implications for financial stability on the role of asset correlations in exacerbating severity and likelihood of systemic crises (Schmieder, 2013; Taleb et al. 2012; Dullmann et al. 2008; and Lopez 2004).

Finally, our estimates on the likelihood and severity of tail events also shed light on the economic and financial impact of the Covid-19 pandemic, thus contributing to most recent research in this area (e.g., Gease and Haldane, 2020; BIS, 2021; Huser et al., 2021; Schrimpf et al., 2020 and Abuzayed et al. 2021).

3. Dataset

The first contribution of our work is the construction of the UK banking system's asset side using a granular approach (Appendix A).[2] The resulting Global Network of UK banks' exposures is composed by six supervisory data sources covering loan, security and derivative exposures as well as secured and unsecured exposures. As shown in Table 1, the Global Network captures £9.4 trillion of gross exposures out of £10.6 trillion of total assets in Q4-2021, roughly 90% of the UK banking system' asset side. The dataset is divided into two main set of exposures. On the one hand, the granular component accounts for 43% of total exposure amounts (£4.1 trillion) and can be split between loan exposures 32.7%, debt and equity security exposures (5.6%), and derivative exposures (5.1%)[3]. On the other hand, when the granular component is not available, we add aggregated exposures by country and sector of the counterparty as residual component which contributes to 57% of the total coverage (or £5.3 trillion)[4].

The total number of entity-to-entity relationships (edges) that we capture are almost 17,556 in Q4 2021 and they are spread across 7,180 counterparties (Nodes_B) and 36 reporting banks (Nodes_L). Overall, the network shows an average counterparty exposure of £1.26 billion (Avg_EXP). Remarkably and consistently with the financial network literature, exposures are

[2] In Appendix A we provide a description of the supervisory datasets, an overview of the methodology to map security and counterparty information, and additional complementary statistics.

[3] Granular exposures are defined as bank to counterparty relationships.

[4] Aggregated exposures are defined as bank to country-sector relationships. For instance, granular exposures towards households are not available, so we complement the dataset using aggregated statistics.

power-law distributed, thereby highlighting a high degree of concentration risk (See Section 4)[5].

Furthermore, by looking at the average path length[6], we notice that the coefficient is quite small averaging around 2.26, implying a quite fast-connected network of relationships, that is, shocks may propagate quickly affecting multiple entities. Nevertheless, the network is not a complete fully-connected network. Its diameter, i.e. the shortest distance between the two most distant nodes in the network is 127, whereas the density parameter[7] is equal to 0.12% emphasizing the UK banking centric perspective of our dataset[8].

Table 1: Summary Global Network Statistics Over Time

Time	TOT	Aggregate	Granular	Loan	Security	Derivative	% LGD	Edges	Nodes_L	Nodes_B	Avg_EXP	Avg Path	Diameter	Density	Power Law
Q1-2018	7926	4634	3292	2348	529	415	44%	16793	36	7361	1.08	2.19	88	0.11	1.55
Q2-2018	8508	4943	3565	2665	515	385	44%	17208	36	7393	1.15	2.20	99	0.11	1.59
Q3-2018	8385	4873	3512	2598	512	402	42%	17327	36	7400	1.13	2.20	99	0.11	1.59
Q4-2018	8508	4895	3614	2673	512	429	42%	17279	36	7390	1.15	2.20	99	0.11	1.55
Q1-2019	8555	5071	3484	2549	510	425	42%	17353	36	7393	1.16	2.22	111	0.11	1.56
Q2-2019	8718	5210	3508	2566	531	411	42%	17757	36	7427	1.17	2.24	133	0.11	1.51
Q3-2019	9066	5552	3514	2547	540	427	41%	17931	36	7426	1.22	2.27	126	0.11	1.57
Q4-2019	8694	5236	3457	2575	490	392	41%	16509	35	6700	1.3	2.24	124	0.12	1.59
Q1-2020	9851	6188	3663	2663	497	503	41%	17117	36	6766	1.46	2.37	134	0.12	1.56
Q2-2020	9750	6005	3745	2742	548	455	42%	17076	36	6839	1.43	2.36	174	0.12	1.55
Q3-2020	9867	5914	3953	2941	524	488	40%	17358	36	6949	1.42	2.34	123	0.12	1.58
Q4-2020	9884	5974	3911	2945	492	474	40%	16469	36	7172	1.38	2.16	123	0.13	1.60
Q1-2021	9577	5607	3970	3020	486	464	40%	17132	36	7346	1.3	2.33	147	0.13	1.57
Q2-2021	9563	5572	3992	3016	502	474	40%	17840	36	7543	1.27	2.33	140	0.13	1.60
Q3-2021	9659	5553	4105	3071	526	508	40%	17985	36	7488	1.29	2.25	163	0.13	1.56
Q4-2021	9444	5344	4100	3087	528	485	40%	17556	36	7180	1.32	2.29	148	0.13	1.59
Average	9122	5411	3712	2750	515	446	41%	17293	36	7236	1.26	2.26	127	0.12	1.57

Note: Values are reported in £ billion for columns (2) to (7). Column "TOT" refers to the total original amount of exposures captured, "Aggregate" refer to the exposure amount mapped on aggregate sector-country counterparty basis , "Granular" refer to the exposure amount mapped with exposure-specific information, of which respectively loan, security and derivative exposures. "% LGD" refers to % of loss-given-default exposures. "Edges" refers to the total number of linkages, "Nodes_L" to the number of reporting banks, "Nodes_B" to the number of counterparties and "AVG_EXP" to the average exposure amount by counterparty. In the end, "Avg_path" refers to the average path length of the network, whereas the "power_law" reports the numeric scalar and exponent of the fitted power-law distribution.

3.1 Sectoral Decomposition

Next, we further investigate the topology of the network by sub-setting the dataset according to the sectoral classification of the counterparty so as to assess the degree of heterogeneity in sector-specific networks as set out in Table 2. In Q4-2021, the most relevant counterparty sector is general governments (GG) capturing 21.8% of total gross exposure amounts. Then it follows exposures to financial corporations (FC) with 21.6%, and, after that, exposures to credit

[5] In this respect, we fit a power law distribution to our network, and we find that the alpha coefficient average around 1.6 across time. For this value of alpha ($\alpha \leq 2$) we can state that our network follows a power-law distribution for which the value of the mean is dominated by the largest exposures in the network (Newman, 2004)[5]. For instance a coefficient of alpha equal to 2.1, which is used to approximate wealth distributions, implies that roughly 80% of total exposure amounts is concentrated in the top 20% exposures.

[6] It is defined as the average number of steps along the shortest paths for all possible pairs of network nodes.

[7] The ratio of the number of edges to the number of possible edges in the network.

[8] We need to recall that the Global Network is composed only by UK banks' exposures, implying that by construction every counterparty which is not a UK bank can't be connected to any another non-UK bank entity.

institutions (CI) with 18%, to the household sectors (HH) with 17.4%, to non-financial corporates (NFC) with 13.3%, and finally to central banks (CB) with 8%.

By comparing the UK banking sector's asset decomposition by sector with one provided for the Euro Area in Montagna et al. (2021), we can notice that the share of UK banks' exposures towards the various sectors is quite aligned. The share of exposures towards FC sector is 21.6% vs 25.4% among the two banking systems, while towards GG is 21.8% vs 20.7%. For the NFC and HH sectors, we have respectively 13.3% instead of 15.1% in the Euro Area, and 17.4% versus 22.8% for the household sector. Finally, interbank exposures (CI) account for slightly more in the UK than in the Euro Area (18% vs 15.5%).

Next, we notice that 46% of the edges (8,001) in the network are directed towards NFCs, which account for roughly 53% (3,800) of the total number of counterparty entities (7,180). On average an exposure towards a NFC is roughly £0.33 billion, which represents the smallest average exposure amount across all sectors. The highest average edge value of £ 15 billion is vis-à-vis CB sector. Next, edge exposures towards households (HH) and governments (GG) show the second and the third highest average edge value, £9.3 and £4.8 billion respectively, although we should note that for the HH sector we deal with exposures aggregated at the country level since we do not have information on granular loan exposures to households. Finally, the average amount per edge towards financial sector entities (FC and CI) tend to be larger, respectively £1 and £2.4 billion, than the average amount per edge towards NFC. Apart from the household sector for which we do not have granular information, we can state that exposures to the public sector on average tend to be larger than exposures toward financial and non-financial corporates[9].

Overall, both the average amount per edge and the average size of counterparty borrowing seem to be aligned, although with some differences to the Euro Area network metrics described in Montagna et al. (2021). In fact, the smallest average edge exposure is also reported in the EA network towards the NFC sector (€ 0.2 billion). FC sector follows with also a very similar average edge exposure of €0.8 billion. Also statistics for exposures to households are quite similar between the two banking sectors, with an average edge exposure of € 6.5 billion. In contrast, the average edge amount vis-à-vis the GG sector is larger for the UK than for the EA, with the latter reporting an average amount per edge of €1 billion. In the end, exposures towards credit institutions is large in the UK than in the EA banking sector, which has an average

[9] We must highlight that there is a high degree of heterogeneity across reporting bank, since the size of exposures is among other variables, a function of the size of a bank's balance sheet.

exposures of €0.4 billion. This second comparison in terms of banking systems highlights that, although the UK and EA banking systems differ in size, with the latter roughly two times bigger than the former, the network of relationships across sectors are quite similar. Banks appear to diversify their exposures similarly across sectors, and across entities belonging to the same sector, independently of their jurisdiction.

Table 2: Sectoral Decomposition Q4-2021

Sector	Total	Aggregate	Granular	Loan	Security	Derivative	% LGD	Edges	Nodes_B	Avg_EXP
GG	2055	265	1789	1441	302	46	46%	1284	432	4.76
FC	2041	1476	565	239	50	276	37%	3809	2019	1.01
CI	1699	1115	584	359	102	123	36%	2938	696	2.44
HH	1645	1645	0	0	0	0	36%	1416	177	9.29
NFC	1252	800	453	366	48	39	40%	8001	3806	0.33
CB	752	43	710	685	25	0	46%	108	50	15.0
Total	9444	5344	4100	3090	528	485	40%	17556	7180	1.32

Note: Values are reported in £ billion for columns (2) to (7). Column "TOT" refers to the total original amount of exposures captured, "Aggregate" refer to the exposure amount mapped on aggregate sector-country counterparty basis , "Granular" refer to the exposure amount mapped with exposure-specific information, of which respectively loan, security and derivative exposures. "% LGD" refers to % of loss-given-default exposures. "Edges" refers to the total number of linkages, "Nodes_B" to the number of counterparties and "AVG_EXP" to the average exposure amount by counterparty in each sector.

3.2 The Global Network

We now proceed with the visualization of our Global Network in its entirety. Figure 1 aims to highlight three main network perspectives of the Global Network, namely relationships by sector, region and community. To produce these network graphs, we assign to each node a colour according to the sector or region the entity belongs to, respectively panel (a) and panel (b), and to the community according to the modularity scores calculated for each entity as in panel (c). Then we colour the edges according to the node's colour receiving the exposure (target node), except for panel (c) for which we assign the colour by type of exposures, blue for loan exposures and red for security exposures. The size of the nodes is proportional to their eigenvector centrality scores in order to highlight the role played by connectedness rather than size in the network.

Firstly, we can observe that the UK Global Network shows a core-periphery structure. On the one hand, the core (*dotted black circle*) is composed by those entities mostly connected to the majority of UK banks, and to which UK banks are mostly exposed in terms of gross exposures. Exposures to core entities are those that overlap across banks' portfolios of securities and loans, exhibiting strong dependency. All sectors are well represented in the core, especially non-financial corporates and credit institutions. On top of that, it is possible also to identify three additional sectoral clusters, respectively financial corporates, governments and households. The core is also well diversified in terms of regional clusters, with a strong presence of entities from every key regions, as seen by the presence of European, Asian, American and British entities. Typically, these entities are well-established multinational corporations or key international public organizations which fund themselves globally, and are quoted on stock exchanges. The core is also well diversified in terms of community structure. In fact the core is a combination of entities belonging to the top-7 communities which roughly account for 98% of the total number of relationships. It is important to highlight that communities are composed of entities amongst which dense connections exists. By contrast, sparse connections exist between entities belonging to different communities. In this respect, we can notice that an overlap does not exist between community structure and regional-sectoral composition.

On the other hand, the periphery of the network is composed by multiple sub-sets of entities, which are clustered together around those UK banks that are exposed to the same set of counterparties. The periphery can be divided into two regions - a region made of weak common dependencies and a region made of exclusive relationships, approximated by the area inside and outside the red dotted circle, respectively. Clusters constituted by weak common dependencies are those in which two or few banks are exposed to the same set of counterparties as seen in A-Type clusters in Panel (a)[10]. These clusters are placed closer to the core than B-Type clusters since they are attracted to the centre by the size of their exposures, the number of relationships, and the number of lender banks involved in the relationships.

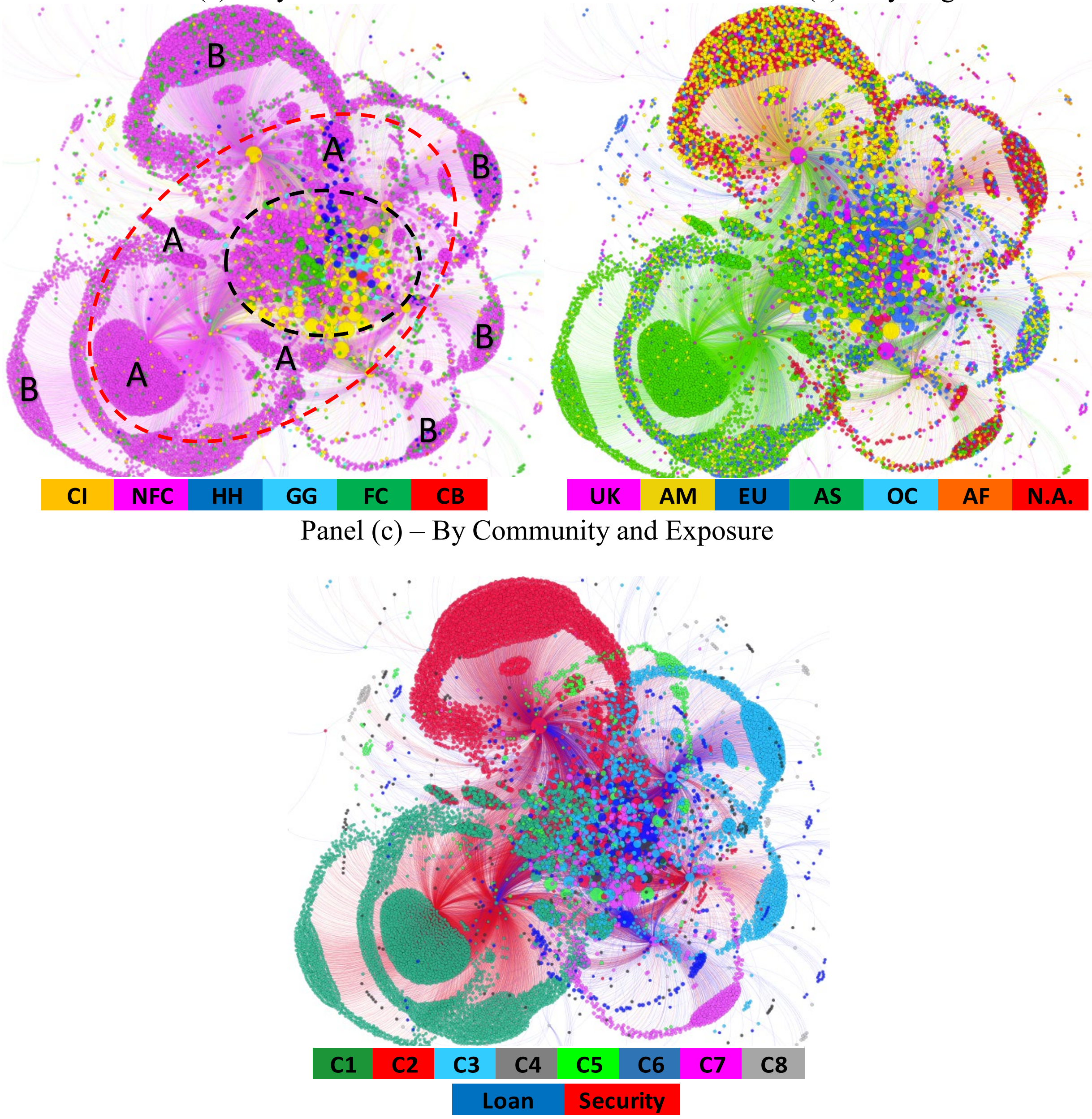
The entities belonging to the same cluster may share the same regional attribute as seen in the variation of colours for these clusters in Panel (b), highlighting common patterns of regional diversification across banks, but it is not exclusive. Exposures to a B-Type cluster are not shared across banks, and do belong to one single bank, thereby representing a bank's exclusive set of counterparties. This would imply that entity-specific shocks to counterparties belonging to B-Type clusters should not directly reverberate through the network if the lender bank is sufficiently capitalized to absorb the losses. Financial contagion may affect other banks only indirectly like via correlations in asset returns or via common macro shocks. On top of this, we need to acknowledge that by construction of the datasets, we are exploiting a UK-banking centric perspective. As such, entities in B-Type clusters and generally any non-UK bank entity might be also linked to UK banks or other corporates via their asset side exposures, which we do not capture in the Global Network.

[10] If those counterparties were exposed to multiple lenders with relative sizable exposures, they would have been placed by the algorithm in the core of the network.

Lastly, we can detect that clusters can be also differentiated by the shares of the type of exposures, as seen in the decomposition between securities and loans in Panel (c). Although security exposures cover less than 20% of total gross exposure amounts, they represent roughly 61% of the total number of edges in the network, 3/4 deriving from debt securities and 1/4 from equity securities.

Figure 1: The Global Network of UK Banks' Exposures

Panel (a) – By Sector Panel (b) – By Region



Note: The total amount of exposures for Q4-2021 is £ 9.4 trillion. The network is built by assigning the eigenvector centrality metrics to the size of the nodes, while the colour of the edges is given by the counterparty node's colour. Blue nodes represent the banking sector, red nodes non-financial corporates, purple nodes the government sector, green nodes the financial corporate sector, and finally the light blue nodes the household sector.

3.3 The Interbank Network

We provide a deep-dive into the topology of the UK interbank network which is constituted by 36 reporting banks and 696 counterparty banks. The interbank network covers roughly £1.7 trillion of exposures in Q4-2021, almost 18% of the total gross exposure amounts as set out in Table 3. Loan exposure account for 61% of granular exposures, whereas security exposures for 18% of total interbank exposures, and derivative exposures for 21%. Nonetheless, 61% of the total number of edges is made of security exposures, of which 78% are bond and 22% equity security exposures, whereas loan exposures account for 24% of the total number of edges in the interbank market, and derivatives for the remaining 15%.

Finally, the average path length coefficient suggests that not all banks are directly connected to each other, with an average of number of steps equal to 2.3. The interbank network is also very sparse, although more dense than the complete network. Among all possible connections, only 0.5% of them are present. Lastly, we see that a small share of interbank exposures accounts for a large share of the total interbank exposure amounts since as is also the case for the complete network, we are dealing with a power-law distribution (alpha coefficient < 2).

Table 3: Interbank Network Characteristics

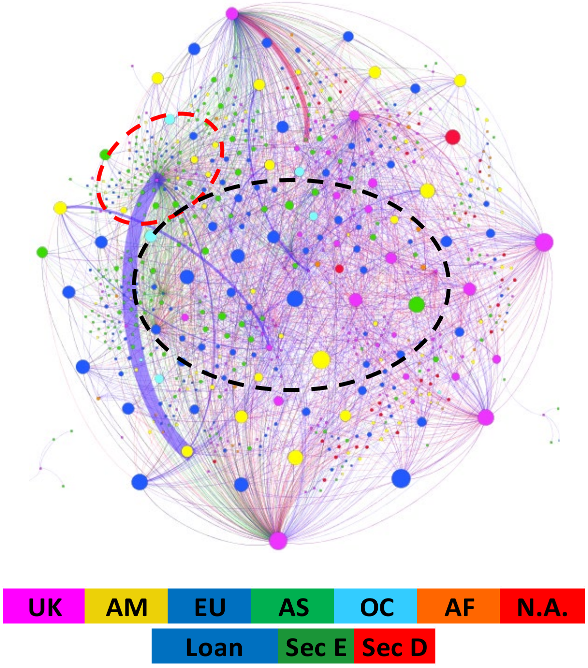
Time	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021	Average
Total	1473	1572	1542	1566	1615	1713	1893	1686	2182	2050	2138	2127	1925	1827	1797	1699	1800
Aggregate	987	1087	1047	1046	1119	1223	1381	1201	1627	1534	1507	1498	1268	1236	1205	1115	1255
Granular	486	485	495	519	496	491	512	485	555	515	631	629	656	590	592	584	545
Loan	288	295	305	318	293	288	306	294	329	295	403	414	449	379	373	359	337
Security	88	86	86	86	85	93	98	93	100	108	109	89	85	94	101	102	94
Derivative	110	104	104	115	118	110	108	98	126	112	119	126	122	117	118	123	114
% LGD	46%	46%	43%	44%	44%	43%	43%	43%	43%	44%	38%	40%	38%	38%	38%	36%	42%
Edges	2782	2845	2855	2848	2892	3014	3015	2907	3082	2984	3039	2807	2899	2949	2992	2938	2928
Nodes	713	704	703	702	703	725	724	686	689	676	703	712	709	720	718	696	705
Avg_EXP	2.1	2.2	2.2	2.2	2.3	2.4	2.6	2.5	3.2	3.0	3.0	3.0	2.7	2.5	2.5	2.4	2.6
Avg Path	2.2	2.2	2.2	2.2	2.2	2.3	2.3	2.2	2.4	2.4	2.4	2.2	2.4	2.3	2.2	2.3	2.3
Diameter	7	7	7	7	7	8	7	7	7	11	6	5	4	5	4	5	6.5
Density	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.4
Power Law	1.84	1.76	1.76	1.79	1.77	1.73	1.71	1.71	1.70	1.84	2.01	1.89	2.29	2.02	2.22	2.01	1.88

Note: Values are reported in £ billion for rows (1) to (6). Row "Total" refers to the total original gross amount of exposures, "Aggregate" refer to the exposure amount mapped on aggregate sector-country counterparty basis, "Granular" refer to the exposure amount mapped with exposure-specific information, of which respectively loan, security and derivative exposures. "% LGD" refers to % of loss-given-default exposures. "Edges" refers to the total number of linkages, "Nodes" to the number of counterparties and "AVG_EXP" to the average exposure amount by counterparty. In the end, "Avg_path" refers to the average path length of the network, whereas the "Power_law" reports the numeric scalar and exponent of the fitted power-law distribution.

By looking at the topology of the network depicted in Figure 2, we bring further insights on the UK interbank network structure. In order to highlight these features, we assign the nodes and edges with colours according to their region and type of exposures, respectively. The size of the nodes is proportional to the eigenvector centrality score, whereas the size of exposures is equal to the gross original exposure amount. We notice that the most central institutions captured by their size are not only UK entities, but also European and American banks, corroborating a strong degree of openness and internationalization of the UK banking sector.

Moreover, Asian banks tend to be less central overall in the UK interbank network in comparison with their share of exposures in the Global Network. The topology of the interbank network also displays a core-periphery structure, the core highlighted by a black dotted circle. The entities belonging to the core are those that are commonly exposed across all UK banks. UK banks also tend to create their own periphery by building their own communities of interbank relationships highlighted for instance by the red dotted circle. We need to acknowledge that our interbank network graph only highlights relationships from UK banks, whereas many other asset-side relationships may exist from non-UK banks towards UK banks as well as non-UK banks, although they are not displayed given our UK-centric data coverage[11]. Lastly, we highlight the composition of the relationships by exploiting the type of exposure, blue for loans, red for debt securities, and green for equities. We notice that, consistent with what we previously stated, the graphical representation highlights a more relevant presence of security exposures over loan exposures, although loan exposures tend to be more sizeable. In this respect, some banks tend to diversify across type of exposures, while others tend to privilege loan and debt security exposures. Only few banks show remarkable interbank equity exposures.

Figure 2: The UK Interbank Network by Country and Type of Exposure



Note: The total amount of exposures for Q3-2021 is £ 1293 billion. The network is built by assigning the eigenvector centrality metrics to the size of the nodes. The colour of the nodes is attributed by the geographical locations of the entities, while the colour of the edges is given by the type of exposure, respectively blue for loan exposures, green for security equity exposures, and red for debt security exposures.

[11] Community-based relationships are important for mitigating liquidity-funding risks. As found in Allen et al. (2020) banks in a community on average have lower centrality of interbank borrowing as expected, nevertheless being in a community can mitigate the negative effect of lacking trust in obtaining interbank funding.

4. Degree of Concentration Risk and Interconnectedness

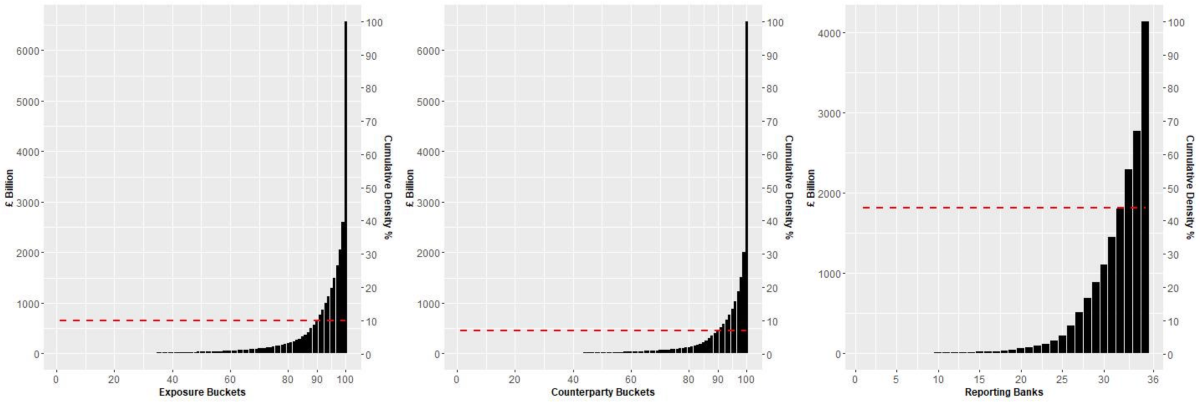
In this section we provide stylized facts on the degree of concentration risk and interconnectedness of UK banks' granular exposures in order to highlight the methodological relevance of these network features in the measurement of tail-risk.

4.1 Concentration Risk

First, we subset the network selecting only those exposures that are granular, thereby omitting aggregate exposures. We are left with roughly £4.4 trillion or 43% of the total coverage. Next, we rank exposures by size and we group them into 100 equal buckets (1% of the total number of exposures). Hence, we construct a cumulative discrete density function. Panel (a) of Figure 3 shows that roughly 60% of the total gross exposure amounts belongs to the 99th percentile, and that the 90th percentile captures roughly 90% of the total coverage. This first result highlights the degree of concentration of UK banks' assets in a small share of large exposures. This is relevant from a modelling perspective since a negative shock to one of these exposures (degree of concentration risk) remarkably affects the stability of the UK banking system. Hence, the distribution of shocks affecting banks' exposures is non-neutral, and determine the shape of the UK banks' loss distribution.

Figure 3: Cumulative Density by Exposure, Counterparty and Reporting bank in Q3-2021

Panel (a) –by Exposure   Panel (b) – by Counterparty   Panel (c) –by Reporting Bank



Note: the red line refers to the 90th percentile.

Moving to Panel (b), we provide a discrete cumulative density of total exposure amounts by counterparty constructed in the same way as presented in Panel (a). In this respect, this cumulative density is even more right-tailed than the exposure-based one, with the 99th percentile capturing roughly 70% of total gross exposure amounts, and the 90th percentile covering 93% of the total[12]. This result corroborates and complements the previous one, emphasizing more strongly the relevance of modelling granularly the distribution of shocks, also at a counterparty level. The failure or distress of certain counterparty entities may put in

[12] Those counterparty entities may be considered as too-big-too fail.

jeopardy the UK banking system' financial stability. Similarly to Gabaix (2011), negative idiosyncratic shocks to the top-1% counterparties may endanger the solvency position of UK banks. Also from this counterparty perspective, the distribution of shocks is non-neutral. Lastly, Panel (c) reports the cumulative density by reporting bank. In this case, we do not group them into buckets since the number of reporting banks is much smaller, whereas each bucket now represents one reporting bank. As seen in Panels (a) and (b), the cumulative density in Panel (c) is also strongly right-tailed, with the top-4 largest banks capturing roughly 55% of total UK banking system's assets. From this lending-side perspective, we may say that also the provision of credit jointly with the portfolio holdings of financial assets are in the hands of few large players. Hence, a negative shock to this set of entities therefore may trigger contagion spillovers across the whole interbank and non-interbank network as demonstrated by multiple studies in the network and financial contagion literature (Covi et al. 2021; Cont and Schanning 2017; Kok and Montagna 2013). Overall, this set of stylized facts highlights the importance of modelling the distribution and transmission of shocks on a granular level since there is high level of concentration risk which can only be captured by modelling entity-to-entity relationships.

4.2 Direct and Indirect Interconnectedness

To further shed light on the critical role played by certain entities in the UK Global Network, we provide entity-specific statistics on their degree of concentration risk, direct and indirect connectedness both from the reporting and counterparty side perspectives.

For the reporting side we compute three main metrics (Equation 1a, 1b, 1c), namely: the concentration of a bank's portfolio as the share of exposure amounts for the 90th percentile over total exposures by reporting bank i ($Conc_i$), the degree of connectedness of each reporting bank i ($Conn_i$) as the number of times bank i appears on the counterparty side[13], and the degree of overlapping portfolios calculated as the summation of the connectedness coefficients ($Conn_i$) across all counterparties j belonging to the portfolio of bank i (OP_i).

$$Conc_i = \frac{\sum_{90\%}^{K} Exp_{i,k}}{\sum_j Exp_{i,j}} \quad (1a); \qquad Conn_i = \sum_i N_i \quad (1b); \qquad OP_i = \sum_j^J Conn_{i,j} \quad (1c)$$

Next, we rank banks by clustering them into colour buckets according to the weighted average of their standardized metrics, with blue for the tier-1 bucket, with green for the tier-2 bucket, and with white for the tier-3 bucket. Panel (a) of Figure 4 plots these metrics into a

[13] We compute this coefficient only one time for all counterparties, and reporting banks do appear on the counterparty side.

three-dimensional graph, in which the X-axis represents concentration risk, the Y-axis Connectedness, and the Z-axis Overlapping Portfolios.

In this respect, focusing on tier-1 banks we identify four entities for illustrative purposes as described in Panel (a) of Table 4. Type-A entities have been placed in the North-Central region (NC), and can be considered the most relevant from a systemic risk monitoring perspective. This set of entities show a very high concentration risk in its portfolio of exposures (its top 10% of exposures represent roughly 90% of its total exposures), meaning it is subject to high idiosyncratic risk. Type-A banks are also very well connected to the other reporting banks (roughly more than 50 banks), meaning that in case of distress, it is likely to trigger direct contagion in the interbank market. Moreover, its portfolio of exposures strongly overlaps with other banks' portfolios, meaning that it may also be subject to the trigger of fire-sales dynamics affecting the whole interbank network indirectly.

By contrast, Type-B entities are located in the North-West region (NW), and can be considered the second most relevant from a systemic risk monitoring perspective. They are prone to a high degree of idiosyncratic risk (high value on the X-axis) and in case of distress they will likely spread risk directly to its interbank peers. However, its portfolio of counterparties is well diversified, and so it will be less susceptible to indirect contagion. Next, Type-C entities in the South-Central region (SC) are the third most relevant in terms of risk-monitoring. In this respect, Type-C entities are subject to a high degree of idiosyncratic risk (X-axis) and indirect risk via overlapping portfolios (Z-axis), nevertheless they won't spread direct contagion in the interbank market (Y-Axis). Lastly, Type-D entities which are located in the South-West region (SW) can be susceptible to idiosyncratic shocks, but show low levels of both direct and indirect contagion spillovers. Finally, it is important to notice that no reporting banks show low levels of concentration risk (X-axis) and at the same time a high degree of both direct and indirect contagion (Y and Z axis).

Table 4: Risk Monitoring Classification

Panel (a)

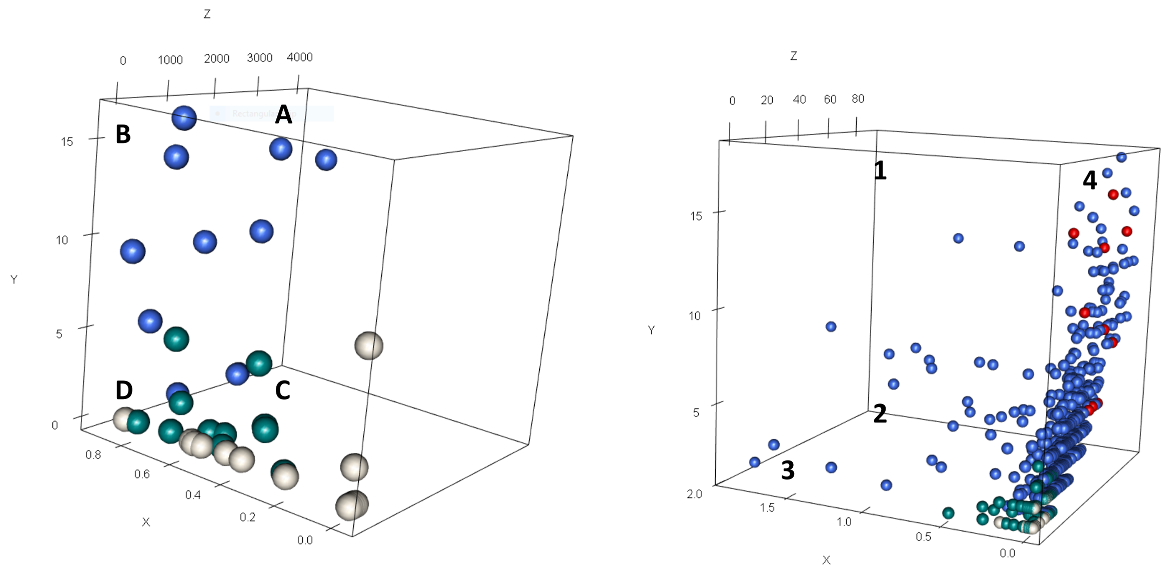
Reporting Entities	A	B	C	D
Idiosyncratic Risk	H	H	H	H
Direct Contagion	H	H	L	L
Indirect Contagior	H	L	H	L
REGION	NC	NW	SC	SW
SYSTEMIC RISK	H	M	M	L

Panel (b)

Counterparty Entities	1	2	3	4
Idiosyncratic Risk	H	H	H	L
Direct Contagion	H	L	L	H
Indirect Contagion	H	H	L	H
REGION	NC	SC	SW	NE
SYSTEMIC RISK	H	M	L	L

Note: "H" stands for high, and "L" for low, and "M" for medium risk. "NC" refers to north-central region, "NW" to north-west, "NE" to north-east, whereas "SC" refers to south-central, and "SW" to south-west.

Figure 4: Concentration, Connectedness, and Overlapping Portfolios
Panel (a) - Reporting Banks Panel (b) – Counterparties



Note: For panel (a) X-axis refers to Concentration, Y-axis refers to Connectedness, and Z-axis refers to Overlapping Portfolio. For panel (b) X-axis refers to Concentration risk measured by the share of exposure amounts by counterparty in % of total exposure amounts in the system, Y-axis refers to Connectedness that is number of banks connected to the entity, and Z-axis refers to Indirect Contagion, first and second layers of connected entities. Red dots highlight reporting banks.

By moving to the analysis of the counterparty side, that is Panel (b) of Figure 4, we compute three main metrics which slightly differ from those previously reported (Equation 2a, 2b, 2c)[14]. Hence, we compute first the concentration risk of a counterparty as the share of gross exposures of counterparty j over total gross exposure amounts ($Conc_j$) divided by $Conn_j$ thereby capturing the average size of exposures. We then compute the degree of connectedness of each counterparty entity j ($Conn_j$)[15], and finally the level of indirect contagion approximated by the summation of the connectedness coefficients ($Conn_j$) across all reporting banks connected to the counterparty j. The colour bucketing is classified as discussed previously with the addition of red dots which identify the UK banks which appear as counterparties.

$$Conc_j = \frac{\frac{\sum_j Exp_{i,j}}{\sum_i Exp_{i,j}}}{Conn_j} \quad (2a); \qquad Conn_j = \sum_j N_j \quad (2b); \qquad IC_j = \sum_i^I Conn_{j,i} \quad (2c)$$

In this respect, focusing on tier-1 counterparties, we identify four types of entities for illustrative purposes as described in Panel (b) of Table 4. These entities are the most relevant from a systemic risk monitoring perspective since they are located in the North-Central region (NC), thereby representing a high degree of idiosyncratic risk (X-axis) for multiple reporting banks (Y-axis), which in turn are very well connected to other banks in the interbank network

[14] We eliminate few outliers on the X-axis for the graphical representation.
[15] This metric is the same as the one computed for the reporting side, that is, the number of reporting banks (i) counterparty j is connected to.

(Z-axis). Type-2 entities which are located in the South-Central region (SC), represent a high degree of idiosyncratic risk (X-axis), but a low level of connectedness since they are connected to only a small fraction of UK banks (Y-axis), though these banks are central in the interbank market (Z-axis). These entities can be considered as medium-risk providers for the system. Next, in the South-West region we find Type-3 entities which are low-risk providers for the system. These entities exhibit a high degree of idiosyncratic risk for their reporting banks (X-axis), but at the same time they are connected to only a very small set of banks (Y-axis) that are not very central in the interbank network (Z-axis). Finally, Type-4 entities are located in the North-East region (NE) and represents a low-risk for the system. This group of counterparties mostly comprises banks (red dots) which represent a low level of idiosyncratic risk for their bank peers. Nevertheless, they are connected with many peers (Y-axis), which in turn are very well connected within the interbank market (Z-axis).

Overall, we can state that degree of concentration risk in the UK banking system both from a reporting and counterparty perspective is high. This characteristic makes the UK banking sector vulnerable to idiosyncratic shocks which via direct and indirect connectedness may spread risk across the global interbank network (Figure 2), within and outside the UK banking system. These stylized facts are informative for the modelling approach we should adopt in order to capture the role played by concentration risk and interconnectedness in the determination of the level of systemic risk.

5. Capital at Risk

In this section we aim at quantifying potential losses of the UK banking system and disentangle their composition. To achieve that, we exploit complementary data on PD and LGD parameters by counterparty sector and country as detailed in Appendix A (section 1.5)[16].

In this respect, two main exercises are provided. First, we compute expected one-year ahead loss estimates, also defined as Capital at Risk estimates (CaR)[17]. Second, in order to incorporate the degree of concentration risk and interconnectedness into our loss estimates we move away from an expected loss calculation methodology and we compute Conditional Capital at Risk Estimates (CCaR) by means of stochastic simulations. Hence we compute conditional loss estimates according to the 90th, 97.5th and 99th percentile of the loss distribution and we thus estimate the severity and probability of the "initial shock". The stochastic approach allows us to model scenario uncertainty, assess the severity of stress scenarios and the trigger event in probabilistic terms. In the end, we decompose the results to shed light on the sources of tail risk for the UK banking sector.

5.1 Measuring Capital at Risk

We compute 1-year ahead expected losses using LGD and PD parameters by sector and country provided by each reporting bank. Since counterparty specific PD parameters are not available, we assign to each counterparty the PD parameter by sector and country averaged across all reporting banks' estimates. This approach resembles a pool-IRB approach of counterparty default rates since we use information reported by all UK reporting banks for each counterparty sector and country. This approach ensures robustness since it is an average estimate across several IRB models[18].

Expected losses are computed using the complete set of exposure - both granular and aggregate - and exposure-based information on the share of unsecured exposure amounts[19]. In this respect, we use a standardized approach to loss calculation and we treat all exposures equally, thereby applying the same LGD and PD parameters to loan, security and derivative

[16] Probability default parameters are based on banks' estimates using their internal models, and are calibrated to the long-run average PD of one-year default rates.

[17] Expected losses are covered by banks' provisions.

[18] The ECB in 2019 has approved the use of pool-IRB approaches in order to better measure PD and LGD parameters. This approach is used especially to estimate counterparty default rates for those type of counterparties whose historical default rates are very low and thus difficult to estimate such as for wholesale exposures (See ECB, 2019).

[19] Hence, we estimate losses using net exposures ($Exp_{i,j}$), thereby deducting on an exposure basis the secured exposure amount from the gross exposure amount so as to derive the exposures at default (EAD).

exposures[20]. Hence, we sum for each reporting bank i expected losses computed on a counterparty basis j and we aggregate them across all reporting banks to achieve a measure of Capital at Risk for the UK banking sector (Equation 3).

$$Capital\ at\ Risk \equiv \sum_{i}^{I} \sum_{j}^{J} Exp_{i,j} * LGD_{i,j} * PD_j \qquad (3)$$

$$where{:}\ i\ refers\ to\ the\ reporting\ bank\ and\ j\ to\ counterparty$$

Figure 5 reports the estimated expected loss amounts decomposed by sector and region of the counterparty. We want to emphasize that loss estimates are for one-year ahead since we use 1-year expected probabilities of default[21].

In terms of Capital at Risk (CaR), the UK banking system's loss estimate in Q4-2021 is close to £31 billion (0.33% of total exposure amounts), up by 17% (£27 billion) relative to the pre-crisis period approximated by estimates for Q4-2019. Importantly the time-series is quite stable over time, ranging between £25 and £28 billion between Q1-2018 and Q4-2019. Though, when the Pandemic starts, the time-series closely tracks the build-up of counterparty risks as the result of the COVID-19 crisis. In Q2-2020, at the peak of the Covid-19 Crisis, total expected losses amounted to £ 37 billion, up by 36% from pre-crisis levels.

Looking at the expected loss estimates by sector reported in Panel (a), most of the losses stem from exposures towards the non-bank financial corporate sector (FC) with £16.6 billion (44.9%), followed by non-financial corporates (NFC) almost £11.7 billion (31.6%), and then by the household sector (HH) £6.5 billion (17.5%). Not surprisingly expected loss estimates vis-à-vis governments (GG) and central banks (CB) account for less than 2% (£0.6 billion), the smallest component, although gross exposures towards Governments account for roughly 21.8% of the total coverage.

Similarly, expected losses towards credit institutions account for only 3% of the total (£1.1 billion), although their gross exposure amounts account for almost 18% of the total. This result is due to a very low average probability of default applied to CI counterparty sector, below 0.2% or 20 basis points, that is, a bank defaulting every 500 years. This result emphasizes the relevance of modelling contagion and amplification effects within the interbank market using microstructural models in order to quantify and factor systemic risk and network risk factors into CI's PD calculations. Overall, the most important component is clearly expected losses

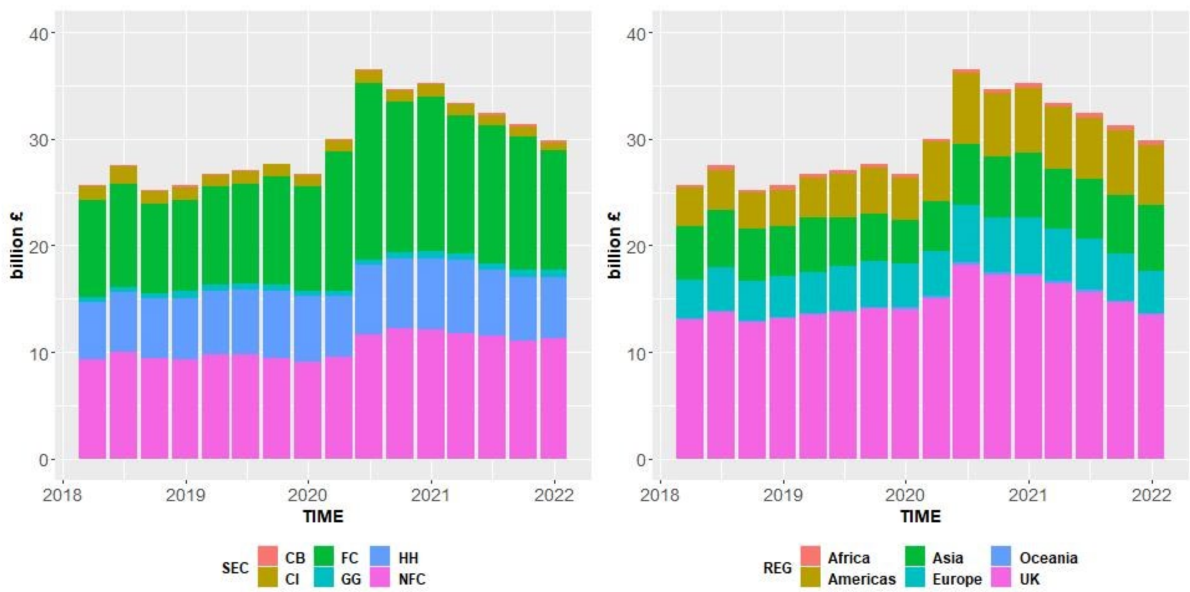
[20] Security exposures concern debt and equity holdings. We do not compute market-risk losses using security-specific parameters such as haircuts and price impacts functions.
[21] For more details on the methodology see BIS (2001). https://www.bis.org/publ/bcbsca05.pdf

stemming from exposures towards the corporate sector, both financial and non-financial, capturing roughly 76.5% of the total (£28.3 billion). This is why, in the credit risk and stress testing literature, most of the estimation effort is spent in accurately estimating losses from counterparties belonging to the corporate sector[22].

Looking at Panel (b) providing the geographical decomposition, loss estimates in Q2-2020 are concentrated within the UK, roughly £18.1 billion (49%), and then equally spread across Americas (17.8%), Asian countries (15.4%) and Europe (14.9%). Finally, expected losses from African countries and Oceania are both below 1%. The geographical loss decomposition corroborates the global nature of UK banks' risk exposure. UK banks tend to import financial shocks from abroad as much as from within the United Kingdom, making the UK financial cycle synchronized with the global economic cycle.

Figure 5: UK Banks' Expected Losses (1-Year Ahead)

Panel (a) – by Sector   Panel (b) – by Geographical Region



5.1.1 Benchmarking

Expected losses increase during the Covid-crisis mainly due to an increase in counterparty risk in the non-financial corporate sector. In Q2-2020 expected losses vis-à-vis the corporate sector (FC+NFC) increase up to £28.3 billion or by 50% between Q4-2019 and Q2-2020.

This finding emphasizes the severity of the Covid-19 pandemic, which necessitates some benchmarking in order to assess the accuracy of the results. Unfortunately, our data availability

[22] In this respect, PD and LGD parameters are estimated using the IRB approach for each counterparty sector and country, exploiting also granular information on type of exposures, for instance differentiating between non-SMEs and SME exposures, and focusing on wholesale credit exposures. Retail exposures and related LGD and PD parameters do not enter into this calculation.

is limited. The Global Network only starts in Q1-2018, thereby not allowing us to compare our results with historical estimates for the 2008 Great Financial Crisis. Nevertheless, we may attempt to compare our estimates with those of other studies, although we should bear in mind that it is difficult to find up-to-date studies with credit loss estimates for UK banks, especially using granular exposure data. In this respect, the BIS Quarterly Review of March 2021 provides credit loss estimates and projections due to the Covid-19 crisis for G7 countries' banking systems using Hardy and Schmieder (2013)' s methodology. This methodology estimates the impact of output on credit loss rates as a non-linear function of both the depth of the recession and its cumulative severity[23]. The report thus provides estimates specifically for the UK, whose increase in cumulative corporate credit losses over the 2020–22 period due to the pandemic amounts to 5.1% of annual GDP (as of 2019), the G7 country most severely affected. This estimate can be translated into roughly £110 billion of credit losses over three years. By calculating our cumulative expected losses over the same period for the corporate sector, our estimates lead to an average cumulative loss of £74 billion[24].

Although the methodologies differ substantially, the BIS report takes a top-down sectoral approach, while ours is a bottom-up granular assessment, results do not remarkably differ. In this respect, our estimates use more up to date data, snapshot ending in Q4-2021 instead of Q4-2020 as in the BIS exercise, thereby incorporating via PD and LGD parameters more up-to-date expectations on the future state of the economy.

5.2 Measuring Conditional Capital at Risk

In this sub-section we move away from the analysis of expected loss estimates, which may be informative on the UK banking sector's average probability of default, and focus on measuring the likelihood and severity of tail events factoring-in the degree of concentration risk and interconnectedness of the UK banking system's network of exposures. Hence, we try to size the impact of those initial shocks that can be considered sizeable as discussed in Acemoglu et al. (2015). In this context, we do not compute amplification and contagion effects stemming

[23] The BIS report calculates the change in credit losses by multiplying the projections of the change in credit loss rates by sectoral credit exposures. In addition to the sectoral credit loss rates, the authors of the report calculate aggregate credit loss rates for each economy based on the aggregate output projections to put the sectoral data into perspective. Based on this approach, the report concludes as follows: "based on our sectoral GDP projections, in a plausible central scenario we find that corporate credit losses during 2020–22 could be equivalent to about three times the pre-crisis level on average across the G7, China and Australia. The additional credit losses emerging from the crisis during the three-year period would cumulate to slightly above 2% of annual GDP or $1 trillion" (BIS, 2021 pp. 68).

[24] We take the quarterly average loss estimates in 2020 (£ 26 billion), in 2021 (£ 24 billion) and we use 2021 estimates (£ 24 billion) as an average projection for 2022.

from interbank financial exposures as the financial network literature does, but instead we treat interbank exposures alike any other counterparty exposure vis-a-vis the other sectors.

From a policy maker perspective this complementary assessment is essential for sizing the sources of potential tail vulnerabilities. It provides a probabilistic assessment of tail-events so as to contextualize the likelihood of their realization at every point in time[25]. Hence, we try to model scenario uncertainty using a stochastic approach to scenario design following Montagna et al. (2021) and Sydow et al. (2021). In contrast to these studies, we compute Conditional Capital at Risk estimates (CCaR), that is, UK banks' losses conditional to the 90^{th}, 97.5^{th} and 99^{th} percentile of the loss distribution.

5.2.1 Tail Events Scenario Design

We split the sample of exposures between granular exposures towards specific counterparty and aggregated exposures towards countries and sectors. For aggregated exposures and granular exposures towards central banks we still compute losses in expectation as the average component, while for granular exposures we compute a distribution of losses, that is, the stochastic component[26]. So the computation of the stochastic component is based on a network of 17,293 edges and 7,236 counterparties which are potentially defaulting entities, covering roughly £3.9 trillion of gross granular exposure amounts (36% of the total coverage)[27]. The stochastic component hence will factor-in the degree of concentration risk and interconnectedness of the UK banking system's network of exposures. The expected loss component instead is calculated upon £4.1 trillion of gross aggregated exposures[28].

Since counterparty specific PD parameters are not available, we assign to each counterparty the PD parameter by sector and country averaged across all reporting banks' estimates. This approach resembles a pool-IRB approach of counterparty default rates since we use information reported by all UK reporting banks for each counterparty sector and country. This approach ensures robustness since it is an average estimate across several IRB models[29].

[25] Top-down stress test exercises are benchmarked on historical macro estimates of extreme events, and so they lose the forward-looking dimension, thereby not providing an estimate for the likelihood of adverse stress scenarios. Using microstructural bottom-up approaches compared to top-down macro models results become a function of exposure and counterparty-specific parameters and ultimately network-specific characteristics.

[26] We don't model stochastically losses vis-à-vis aggregate exposures since it would imply that an entire sector for a country would default, and so all exposures vis-à-vis that sector. Hence we measure losses in expectations for the share of exposures that we can't map granularly. The only exception is the sector central banks (CB) whose losses are model in expectation.

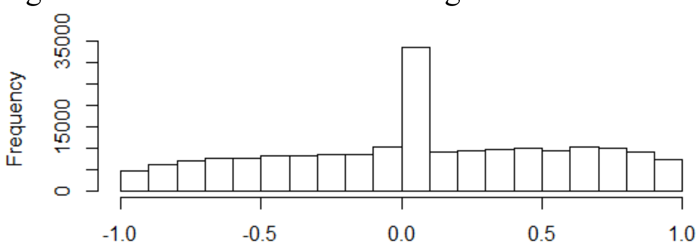
[27] We exclude granular exposures towards central banks from this calculation, which account for £710 billion or 7.5% of total exposure amounts.

[28] Alike the expected loss exercise we use only the unsecured exposure amounts for calculating loss estimates for both the stochastic and average loss components. .

[29] The ECB in 2019 has approved the use of pool-IRB approaches in order to better measure PD and LGD parameters. This approach is used especially to estimate counterparty default rates for those type of counterparties

Next in order to model the level of interdependency in the intersectoral input-output firms' linkages - network origins - as in Acemoglu et al. (2012), we estimate a correlation structure of counterparty defaults (based on reported counterparty PD) by country and sector exploiting a panel dataset covering 134 countries and 4 sectors (NFC, FC, CI, GG) ranging between Q1-2018 and Q4-2021. This leads to an average correlation across countries and sectors for the full sample of 0.045 as reported in Figure 6[30]. Nevertheless there is lot of heterogeneity across sectors and countries, with negative correlation coefficients for some pairs. Hence, we should acknowledge that depending on the time period of the sample we use, although we keep it fixed, the correlation structure should change. The correlation matrix based on counterparty PD by country and sector have been estimated over 16 quarters during which the Covid-19 pandemic took place, leading to a strengthening of correlation across countries and sectors. The average correlation across countries and sectors for the subsample Q1-2018 till Q4-2019 (pre Covid-19) is close to 0.03, whereas for the sub-sample covering the Covid-19 crisis the average correlation coefficient was higher, close to 0.08.

Figure 6: Distribution of the Average Correlation Coefficients between Q1-2018 and Q4-2021



Note: Average across all correlation parameters by country-sector pairs equal to 0.045, thereby approximating median stress events as estimated in Hardy and Schmieder (2013).

Correlation coefficients do change over the business cycle and are affected by macroeconomic and financial conditions. For instance, Hardy and Schmieder (2013) relies on the empirical estimates of Duellmann et al. (2008) who provide estimates on fluctuations of asset correlations under macroeconomic stress for a sample of large Western European corporates between 1997–2003. Conditional to their data sample, asset correlations fluctuated strongly, ranging from 0.04 to 0.16, with a mean at 0.1. Overall, our estimates are lower than Duellmann et al. (2008) due to differences in the sectoral coverage, since the estimates given here are an average across sectors, and also include financial corporates and governments which experience much lower historical probability of defaults. Hence, by sub-setting our

whose historical default rates are very low and thus difficult to estimate such as for wholesale exposures (See ECB, 2019).

[30] In this exercise the correlation is time-invariant. We use the average across all reported quarters. The peak of the distribution coincides with the average correlation coefficient since we fill NAs with the average correlation coefficient.

sample to only non-financial corporates (NFC), we find that the average correlation coefficients for the corporate sector across countries is close to 0.1, and respectively 0.07 before the Covid-19 crisis (Q1-2018 to Q4-2019), and 0.13 during the Covid-19 crisis (Q1-2020 to Q4-2021). Hence, our estimates are aligned with and corroborates Duellmann et al. (2008)'s estimates for the pre-crisis period for the corporate sector.

Moreover, Hardy and Schmieder (2013)'s estimates for asset correlation differ conditional to the degree of stress scenario. For instance, during normal conditions (median), medium stress (90th percentile) and severe stress conditions (97.5th percentile) asset correlation increases respectively from 0.1, to 0.22 and 0.3. In this respect, our average correlation estimates for the corporate sector during the Covid-19 period (0.13) can be classified in between normal and medium stress conditions.[31]. This is due to the type of stressed correlation that the Covid-19 crisis has produced, that is, a heterogeneous stress across countries and sectors. Moreover, we should emphasize that also the timing was different, since not all the countries have been hit simultaneously. This heterogeneous shock which is reflected in change in counterparty PDs may increase the positive correlation for certain country-sector pairs or decrease it for other country-sector pairs. This is consistent with the BIS (2019)'s results on credit loss estimates for the G7 countries' banking systems, since they conclude that the Covid-19 crisis seems not to be as severe as the GFC[32] and closer to a medium stress event[33]. Overall, we can state that our correlation matrix of counterparty default probabilities averaged over the full sample period can be considered a reliable proxy for medium-low stress events, thereby quite conservative[34]. In Section 5.3 we will test our results to this assumption.

We thus produce 20,000 Montecarlo simulations of Bernulli vectors of corporate defaults for 7,236 counterparties by modelling an M-variate distribution with uniform marginals and a Gaussian copula with a covariance matrix characterized by the correlation structure above estimated (Montagna et al. 2021). As reported in Equation (4), each entry ($J_{j,s}$) of the Bernulli vector take value 1 if the counterparty defaulted or value 0 if it did not. Hence, for each simulation/scenario, we compute the stochastic loss component for each reporting bank by

[31] We need to emphasize that the Covid-19 crisis has affected sectors and countries heterogeneously, for this reason the average correlation coefficient did not increase remarkably.

[32] These results are based on the parameters and elasticities provided in Hardy and Schmieder (2013).

[33] In this respect, the study used parameters and elasticities for credit loss estimates calculation that were an average between medium and severe stress events.

[34] We could have calculate a correlation structure for different sub-periods, but this would create problems in comparing stochastic loss estimates across quarters. Hence, we prioritize comparability of results across time. Moreover, as emphasized by Schmieder et al. (2011) fixed IRB correlations based on low PDs can be used as benchmark estimate for supervisory purposes. See also Lopez (2004).

multiplying exposures at default ($Exp_{i,j}$) by the counterparty-specific LGD parameter and by 1 or 0 depending on whether the counterparty in that scenario defaulted or not. We then sum across each reporting bank in order to obtain the stochastic loss component for the UK banking system in each scenario. Finally - Equation (5) - we sum the stochastic loss component of each scenario - $SCaR_{G,s}$ - (derived from the granular exposures network) and the expected loss component - $ECaR_A$ - (derived from the aggregate exposures network) to obtain the loss distribution for the whole UK banking system. In the end by conditioning the resulted loss distribution to the selected percentile x (99th to approximate extreme stress events, 97.5th for severe distress, and 90th for medium stress), we derive our measure of Conditional Capital at Risk (CCaR) as reported in Equation (6).

$$Stochastic\ Capital\ at\ Risk_s \equiv \sum_{i}^{I}\sum_{j}^{J} Exp_{i,j} * LGD_{i,j} * \begin{bmatrix} J_{1,1} & J_{1,2} & J_{1,s} \\ J_{2,1} & J_{2,2} & J_{2,s} \\ J_{j,1} & J_{j,2} & J_{j,s} \end{bmatrix} \tag{4}$$

Where i refers to the reporting bank, j to each counterparty, s to the realization of each single scenario, and $J_{j,s}$ takes value 1 if counterparty j in scenario s defaults and 0 if it doesn't.

$$Loss\ Distribution \equiv SCaR_{G,s} + ECaR_A \tag{5}$$

$$Conditional\ Capital\ at\ Risk = Loss\ Distribution^x \tag{6}$$

Where G refers to the network of granular exposures, A to the network of aggregated exposures, and x to the percentile of the loss distribution.

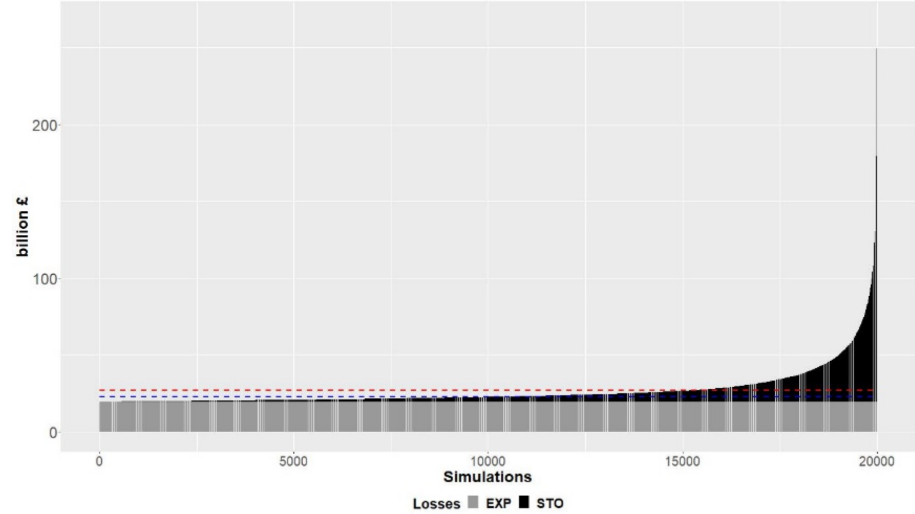
5.2.2 Severity and Likelihood of Tail Events

We provide distributions of total loss estimates (CCaR estimates) for the UK banking system for all quarters (Table 5), decomposing them by the expected loss component and the stochastic loss component. Starting from a visual inspection of Figure 7 which reports the loss distribution for Q4-2019 (pre-crisis period) total losses differ remarkably between the left tail (low severity scenarios) and the right tail (high severity scenarios) of the distribution. As we move closer to the right tail, we can see that the share of stochastic losses increases over the share of expected losses. This evidence shows a high dispersion even in "normal times", with the most extreme scenario being 10 times worse than the average scenario. Losses in the 99th and 97.5th percentile scenarios classified as extreme and severe stress events amount up to respectively £89 billion and £66 billion, roughly 3.3 and 2.4 times more than an average scenario[35]. Hence, modelling

[35] In this respect, the BOE's 2019 stress testing exercise estimates roughly £180 billion of losses from credit and traded risks over a five year period given an adverse scenario which resembles macroeconomic and financial conditions experienced during the GFC. Most of the losses take place in the first two years of the scenario.

and analysing scenario uncertainty is very important to shed light upon the *severity* of tail events.

Figure 7: UK Banks' Stochastic Distribution of Losses in Normal Times (Q4-2019)



Note: grey bars refer to the expected loss component and they are equal by construction across all scenarios. Black bars refer to the stochastic loss component and they are ranked from left to right by severity. Red line refers to the average loss, while the blue line to the median loss.

Nevertheless, we need also to contextualize this value in probabilistic terms. In this respect, the likelihood of experiencing an extreme event in Q4-2019 of £91 billion losses or above the average loss of the 99th percentile during the pre-pandemic period (Q1-2018 till Q4-2019), which can be considered an extreme stress event based on the above definition of Hardy and Schmieder (2013), is low, that is, 296 scenarios over 20,000 simulations, hence a probability of 1.5% (once every 67 years)[36]. However, if we benchmark the probability against experiencing a severe stress event (97.5th percentile), that is, above £65 billion, the probability increases up to 3.9% (once every 26 years). In the end, if we benchmark the probability against experiencing an medium stress event (90th percentile), that is, above £35 billion, the probability increases to 10.4% (once every 10 years).

Table 5: UK Banks' Stochastic Distribution of Losses – Benchmark Case

STATISTICS	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
Mean	26	28	25	26	27	27	28	27	30	37	35	35	33	32	31	31
Median	22	23	21	21	22	23	23	22	24	29	29	30	28	27	26	26
Pct 99%	88	101	86	90	90	89	91	104	117	147	116	114	107	110	108	108
Pct 97.5%	64	73	63	62	63	66	65	76	87	109	86	86	77	77	79	79
Pct 90%	34	37	33	34	35	37	37	36	41	52	49	49	46	45	44	44
Prob 99%	0.9%	1.4%	0.9%	1.0%	1.0%	1.0%	1.0%	1.5%	2.2%	4.1%	2.1%	2.2%	1.7%	1.7%	1.7%	1.7%
Prob 97.5%	2.3%	3.2%	2.2%	2.2%	2.4%	2.5%	2.5%	3.9%	4.7%	6.7%	5.1%	4.9%	3.9%	4.0%	3.9%	3.9%
Prob 90%	9.4%	11.2%	8.5%	8.8%	9.9%	11.4%	11.3%	10.4%	13.2%	21.9%	21.2%	22.6%	19.7%	17.8%	16.6%	16.6%

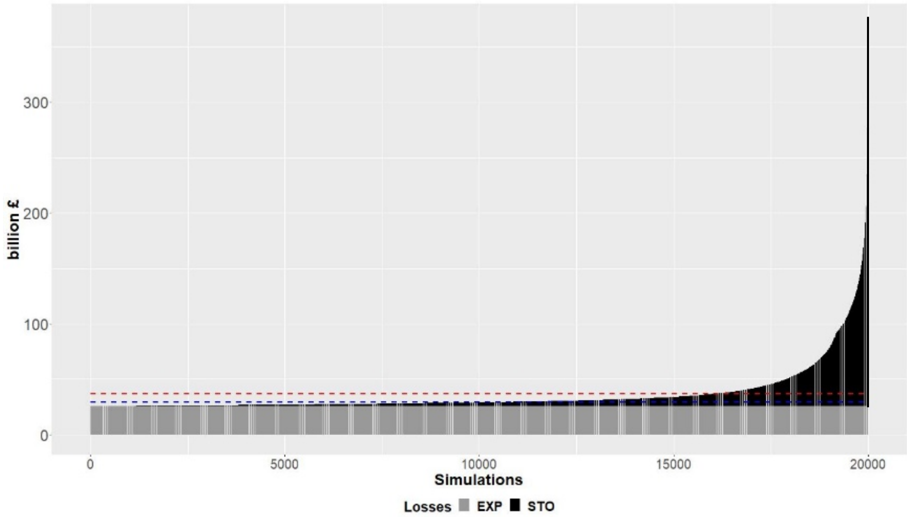
Note: "Mean" refers to average banks' losses, and "Median" refers to the median losses, "Pct 99%", "Pct 97.5%", "Pct 90%" refers respectively to the loss for the 99th, 97.5th and 90th percentile of the loss distribution representing extreme, severe and medium stress events. In the end "Prob 99%", "Prob 97.5%" and "Prob 90%" refer to the probability of extreme, severe and

[36] We compute the average loss over the pre-pandemic period (Q1-2018 to Q4-2019) for conditioning tail-risk to the 90th, 97.5th and 99th percentile since we want to capture deviations relative to a stable period, and smooth out quarterly variations.

medium stress events, respectively with losses higher than £91 billion, £ 67 billion and £35 billion benchmarked on average losses for the same percentiles during Q1-2018 to Q4-2019. The number of simulations equal to 20,000.

However, macroeconomic and financial conditions may change over the business cycle and given unexpected shocks such as the Covid-19 pandemic. In this respect, Figure 8 provides the stochastic distribution of banks' losses calculated in Q2-2020, that is, at the Peak of the Covid-19 crisis. First of all, the mass in the right tail of the distribution remarkably increased, and the stochastic component increases its relevance relative to the expected loss component, thereby increasing the overall severity of stress events. Overall, the severity of extreme events increased remarkably during the Covid-19 pandemic, that is, given a worsening of financial and macroeconomic conditions which are reflected in higher levels of counterparty risk[37].

Figure 8: UK Banks' Stochastic Distribution of Losses during the Covid-19 Crisis (Q2-2020)



Note: grey bars refer to the expected loss component and they are equal by construction across all scenarios. Black bars refer to the stochastic loss component and they are ranked from left to right by severity. Red line refers to the average loss, while the blue line to the median loss. The number of simulations equal to 20,000.

However, not only did the severity increase, but also the likelihood of experiencing extreme events. In order to properly benchmark the likelihood with the results presented for Q4-2019, we keep the threshold to identify extreme stress events constant at £91 billion of losses, that is, the average loss for the 99th percentile between Q1-2018 and Q4-2019. Hence, we try to answer the following question: "what is the probability of experiencing in Q2-2020 - at the peak of the Covid-19 crisis - the same type of 'extreme stress events' as classified during the pre-crisis period?". The answer is 4.1%, as 828 scenarios (over 20000 simulations) show total losses higher than £91 billion in Q2-2020. In Q4-2019 it was 1.5%, moving from 1 severe distress scenario every 67 years, to 1 severe distress scenario every 24 years. However, if we benchmark the probability against experiencing a "severe stress" event (97.5th percentile or £ 65 billion), the probability increases to 6.7%, that is, 1334 scenarios over 20000 simulations (once every

[37] The correlation matrix is fixed across quarters.

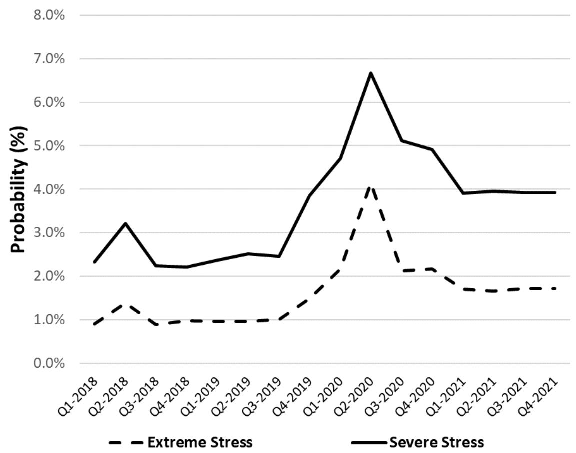
15 years). Overall, we have seen how both the severity and probability of severe stress events increased remarkably during the pandemic given the change in macro and financial conditions.

In the end for monitoring purposes we provide a time-series evolution of probabilities and severity of Tail Events - Figure 9. In this respect, we compute the probabilities and severity for extreme and severe stress events (99th and 97.5th percentiles). The build-up of systemic risk in the UK banking sector approximated by the probability - Panel (a) - and severity - Panel (b) - of extreme and severe stress events by the start of the Covid-19 pandemic is very clear.

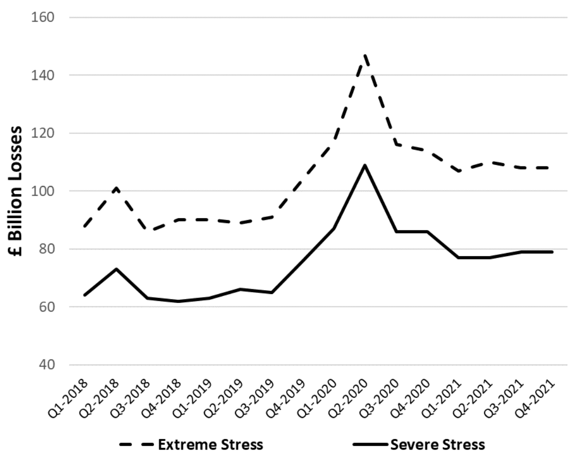
By comparing the increase in the severity and likelihood of tail events with the increase in expected losses (CaR estimates) at the peak of the crisis (Q2-2020), we can see that average risk increase (37%) by less than the increase in tail risk, respectively 310% and 62% for the probability and severity of extreme stress events. This finding highlights that average risk is not a good proxy for tail risk. In Q4-2021, both tail risk measures, although they declined after reaching the peak in Q2-2020, remain still higher than the average pre-crisis levels by respectively 70% (probability) and 20% (severity).

Figure 9: Probabilities and Severity of Stress Events for the UK banking Sector

Panel (a) - Probability



Panel (b) – Severity (CCaR)



Note: Extreme and severe stress events refer respectively to the 99th and 97.5th percentile of the loss distribution. The probability estimates for extreme and severe stress events are computed conditional to losses higher than £91 billion and £67 billion, that is, average losses for the same percentiles during Q1-2018 to Q4-2019. The number of simulations equal to 20,000.

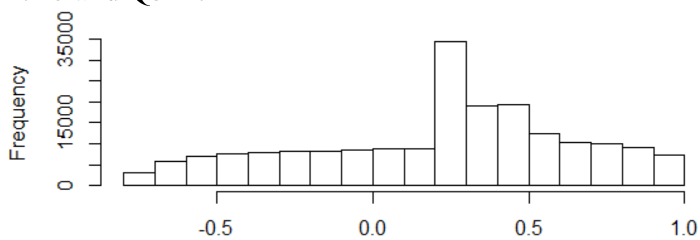
5.3 The Role of Asset Correlation

Asset correlations, as emphasized by Hardy and Schmieder (2013) and by Duellmann et al. (2008), are key inputs for modelling correlated default probabilities in solvency stress test exercises. Usually in standard solvency stress test exercises, the focus is placed upon the estimation of losses in expectation conditional to an adverse scenario, for which the role of correlated shocks and correlated defaults is neglected. As we can see from the results in the previous section, average loss estimates are not affected by the correlation structure and size. However, when we study tail events correlations play a key role in determining their severity and likelihood. Hence we assume that there is a strengthening in the interdependency of input-output relationships of banks' counterparties[38].

In the previous exercise we were using a correlation matrix estimated over the period Q1-2018 and Q4-2021. The average correlation parameter across all countries and sectors was 0.045, and 0.1 for specifically the corporate sector. This was a conservative assumption, since during crisis times, asset correlations tend to strengthen, respectively close to 0.22 for medium stress events and close to 0.3 for severe stress events.

Hence, in this exercise we want to test the sensitivity of our results to an increase in the correlation structure of shocks. To achieve that, first we estimate the correlation matrix over the crisis period Q1-2020 to Q4-2021 in order to maintain the correlation structure between country-sector pairs as it was during the Covid-19 crisis. This leads us to an average correlation coefficient close to 0.08, still too low compared to Hardy and Schmieder (2013)'s estimate for medium stress events. Hence, we reduce negative correlation coefficients and increase positive correlation coefficients by a constant equal to 0.2 in order to make the mean and peak of the distribution coincides with an average correlation coefficient of 0.22 as estimated by Hardy and Schmieder (2013) - Figure 10.

Figure 10: Distribution of Stress Correlation Coefficients over the entire sample period Q1-2018 and Q3-2021



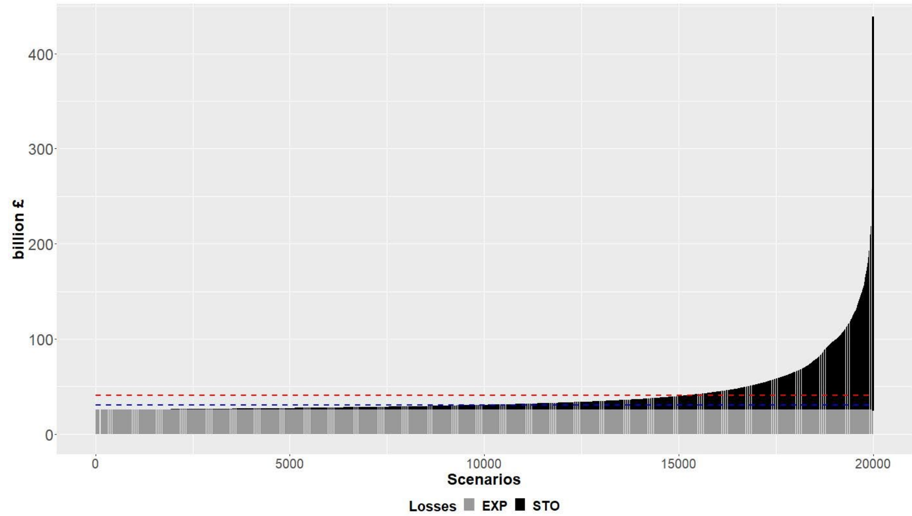
Note: Average across all correlation parameters by country-sector pairs equal to 0.22, thereby approximating medium stress events as estimated in Hardy and Schmieder (2013).

By inspecting Figure 11 and Table 6, we notice that the severity of extreme events further increase relative to the previous baseline case. The right tail of the distribution gets fatter and taller. The most extreme scenario now is above 400 billion of losses and tail events tend to account for an even larger share of total cumulated losses. Hence severity of medium, severe and extreme stress events approximated by CCaR estimates increase as correlation in firms' PDs increases.

---

[38] For instance this may be due to a stronger synchronization of counterparties' revenues across sectors and economic activities (supply chain relationships like just-in-time manufacturing) or due to a synchronization of business cycles across countries. More in general common sectoral shocks like the Covid-19 Pandemic or like climate-related transition risks may generate a strengthening in asset correlations as well as counterparty defaults across sectors and countries.

Figure 11: UK Banks' Stochastic Distribution of Losses during the Covid-19 Crisis (Q2-2020) conditional to a Stress Correlation Structure in Counterparty Defaults



Note: grey bars refer to the expected loss component and they are equal by construction across all scenarios. Black bars refer to the stochastic loss component and they are ranked from left to right by severity. Red line refers to the average loss, while the blue line to the median loss.

To compare the estimated likelihood of extreme events with the previous exercise, we condition the set of stress scenarios to those scenarios with an average loss above £91 billion for the 99th percentile, £67 billion for the 97.5th percentile, and £35 billion for the 90th percentile. Given this stress correlation structure, the probability of experiencing extreme stress events (99th percentile) in Q2-2020 increase to 6% from 4.1%, that is, we estimate 1206 scenarios with total losses larger than £91 billion (over 20,000 simulations). For severe stress events, the probability increases to 10.2%, that is, 2036 scenarios over 20,000 simulations with losses above £67 billion.

Table 6: UK Banks' Stochastic Distribution of Losses based on a Stressed Correlation Matrix

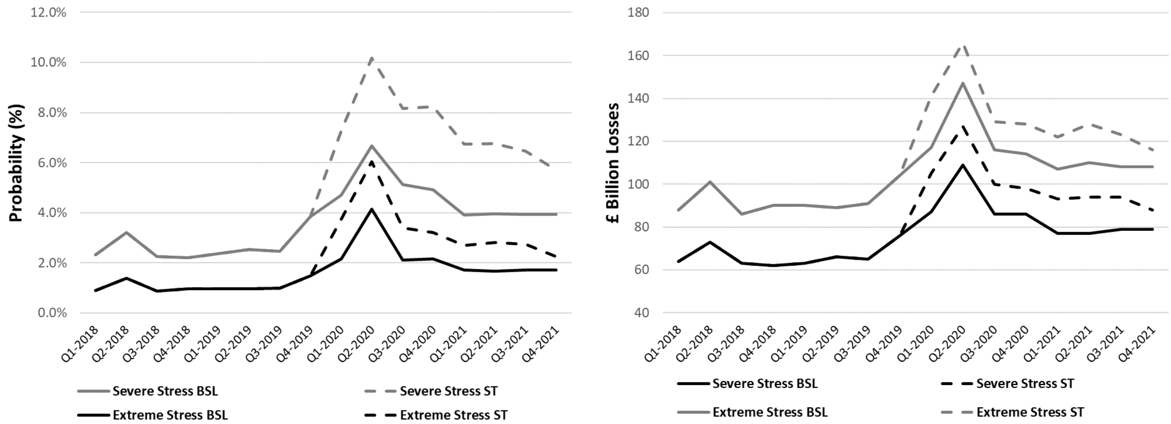
STATISTICS	Medium-Stress Correlation Structure								Baseline Correlation Structure							
	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
Avg_tot	34	41	39	39	37	36	35	33	30	37	35	35	33	32	31	31
Median	26	31	31	31	30	29	28	27	24	29	29	30	28	27	26	26
Pct 99%	141	166	129	128	122	128	123	116	117	147	116	114	107	110	108	108
Pct 97.5%	105	127	100	98	93	94	94	88	87	109	86	86	77	77	79	79
Pct 90%	54	66	60	60	56	56	54	51	41	52	49	49	46	45	44	44
Prob 99%	3.7%	6.0%	3.4%	3.2%	2.7%	2.8%	2.7%	2.3%	2.2%	4.1%	2.1%	2.2%	1.7%	1.7%	1.7%	1.7%
Prob 97.5%	7.3%	10.2%	8.2%	8.2%	6.7%	6.8%	6.5%	5.7%	4.7%	6.7%	5.1%	4.9%	3.9%	4.0%	3.9%	3.9%
Prob 90%	22%	34%	33%	35%	31%	28%	26%	23%	13.2%	21.9%	21.2%	22.6%	19.7%	17.8%	16.6%	16.6%

Note: "Mean" refers to average banks' losses, and "Median" refers to the median losses, "Pct 99%", "Pct 97.5%", "Pct 90%" refers respectively to the loss for the 99th, 97.5th and 90th percentile of the loss distribution representing extreme, severe and medium stress events. In the end "Prob 99%", "Prob 97.5%" and "Prob 90%" refer to the probability of extreme, severe and medium stress events, respectively with losses higher than £91 billion, £ 67 billion and £35 billion benchmarked on average losses for the same percentiles during Q1-2018 to Q4-2019. The number of simulations equal to 20,000.

In the end, Figure 12 reports the probability and severity of extreme and severe stress events over time according to a stress and a baseline correlation structure (straight versus dotted lines). We can note from the visual inspection that the probability and severity of stress events in each percentile of the distribution shifted up starting in Q1-2020[39]. In Q4-2021 both probability and severity of stress events (ST) still remain well above the pre-crisis levels, even more if counterparty defaults are highly correlated. We want to emphasize that these estimates are some kind conservative since we use a stress correlation structure that equal a medium stress scenario (not a severe stress) as classified by Hardy and Schmieder (2013). Furthermore we want to emphasize that expected losses do not change at all, since correlation does not matters in expectation, but only affects the shape of the loss distribution. To conclude, we want to emphasize that PD and LGD parameters in this exercise remain equal to those in the baseline case, and variations are due to only change in the correlation matrix.

Figure 12: Comparison of Probability and Severity of Stress Events for the UK Banking Sector conditional to a Stress and Baseline Correlation Structure

Panel (a) - Probability Panel (b) – Severity (CCaR)



Note: ST refers to results estimated conditional to a stress correlation structure (average equal to 0.22) approximating medium stress events. Whereas BSL refers to baseline results. Extreme and severe stress events refer respectively to the 99th and 97.5th percentile of the loss distribution. The probability estimates for extreme and severe stress events are computed conditional to losses higher than £91 billion and £67 billion, that is, average losses for the same percentiles during Q1-2018 to Q4-2019. The number of simulations equal to 20,000.

5.4 Conditional Capital at Risk Decomposition

In this subsection, we decompose UK banks' CCaR estimates into their sectoral and regional contribution in order to identify the sources of tail events. We focus the analysis on extreme stress scenarios, that is, those scenarios in the 99th percentile of the loss distribution. Results are provided for Q4-2021 in order to capture the latest developments, although Appendix B provides the decomposition over the entire sample period. Table 7 provides a breakdown of average losses for extreme stress scenarios by sector. The decomposition is provided only for the stochastic component, since the decomposition of the expected loss component was provided in Section 5.1.

[39] The stress correlation structure (ST) has been used only for the estimation of the crisis period between Q1-2020 and Q4-2021.

The most important sector is non-bank financial corporates (FC) with a contribution (loss share) of 50.6% to total loss in extreme stress scenarios. Importantly the FC sector also shows a loss ratio of 10.6% as share of FC granular gross exposures. This estimate approximates the share of FC exposures that potentially is subject to default conditional to an extreme stress scenario. Then there follows non-financial corporates (NFC) which account for only 32.6% of total average losses and governments (GG) with 15.8%. As we can see the banking sector (CI) is a small contributor to total losses in extreme stress scenarios, although their total granular exposures combined account for 17% of total granular exposures. This is due to multiple reasons. First probabilities of defaults for counterparty sector CI are very low, ranging between 0.1% and 0.3% across developed regions. Moreover, the correlation structure across countries is three time stronger for the corporate sector (FC and NFC) rather than the CI sector. This implies that especially in the 99th percentile of scenarios - the extreme ones - clusters of counterparty defaults will take place within the sector with the higher correlation structure such as the corporate sector and with higher PDs. In the end, we need to emphasize that our methodology does not capture amplification and contagion effects taking place within the financial system. Contagion and amplification mechanisms such as asset fire-sales will strengthen correlated defaults within and among the CI and FC sectors as well as their loss contribution to severe stress events as shown in Montagna et al. (2021). This methodological extension goes beyond the scope of this paper, and is discussed in the Conclusion.

Finally, we should emphasize that the contribution of FCs and NFCs could potentially be much higher since we capture respectively 28% and 36%of total exposures with granular information[40]. Overall, the total loss ratio across all sectors for extreme stress scenarios is equal to 3.5% of total granular exposures and the loss share across sectors is stable across the entire sample period as reported in Appendix B.

Table 7: Sectoral Decomposition of CCaR Estimates (Q4-2021)

Sectors	FC	NFC	GG	CI	Total
Exposures (£bn)	565	453	1789	584	3391
AVG Loss (£bn)	61	39	19	1	120
Loss Ratio (%)	10.7%	8.6%	1.1%	0.2%	3.5%
Loss Share (%)	50.6%	32.6%	15.8%	1.0%	/

Next, Table 8 provides a breakdown of tail losses for extreme stress scenarios by geographical location. The highest contributor in Q4-2021 is the UK with a share of 47%, whereas Americas are the second highest contributor with 22.4%, followed by Europe with 20.8% and Asian

[40] The household sector is missing since we don't have granular exposures for this sector.

countries with 8.7%. To what may concern the loss ratio, the UK shows the second highest among all geographical regions (5.4%) after Africa (6.6%). Then follows "Americas" with a value of 4%, Europe with 3.1%, and Asia with 1.3%.

Overall, by looking at the time series reported in Appendix B, we can see how the sectoral contribution of tail risk has remained quite stable with only the FC sector experiencing an increased share during the peak of the Pandemic, for then in Q4-2021 returning back to equilibrium. Contrary the geographical contribution to tail risk seems to have increased in Americas and Asia in 2021 and reduced in UK and Europe. This result highlights how the Pandemic has heterogeneous economic consequences across sectors but also across regions and countries. The Covid-19 crisis has affected the sources of systemic risk. Overall, we can state that tail events affecting the UK banking sector in Q4-2021 are more likely to stem from regions outside the UK (53%), corroborating the fact that the UK banking sector is clearly exposed internationally as much as domestically as a small open economy suggests.

Table 8: Regional Decomposition of CCaR estimates (Q4-2021)

Regions	UK	Americas	Europe	Asia	Africa	Oceania	Total
Exposures (£bn)	1035	671	805	823	15	30	3381
AVG Loss (£bn)	56.4	26.9	24.9	10.4	1.0	0.3	120
Loss Ratio (%)	5.4%	4.0%	3.1%	1.3%	6.6%	0.9%	3.5%
Loss Share (%)	47.0%	22.4%	20.8%	8.7%	0.8%	0.2%	/

5.5 Sensitivity Analysis

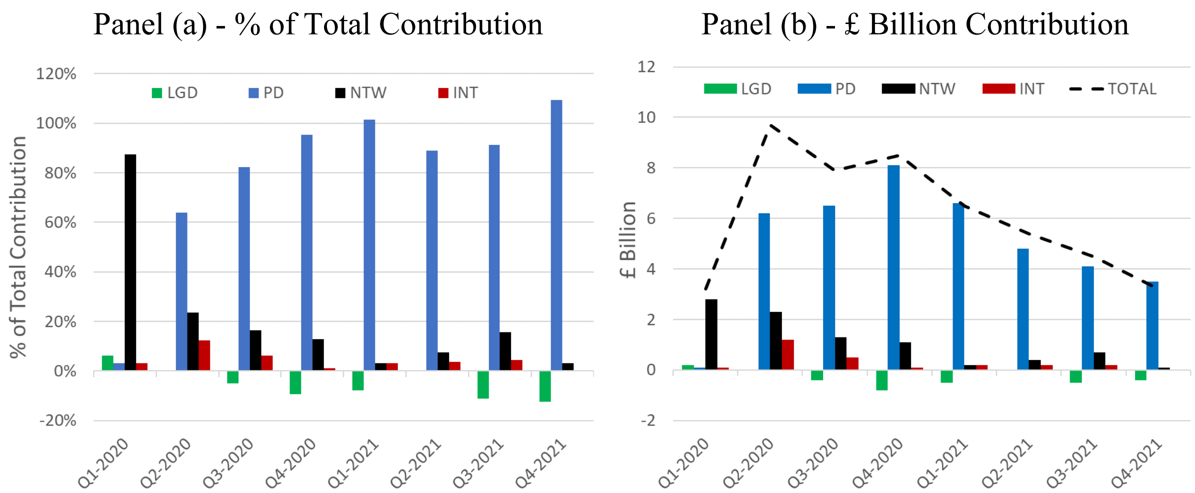
In this section, we provide evidence of the main risk drivers for average and tail risks by mean of counterfactual exercises. We perform three main counterfactuals to compute the loss contribution due to variations over-time in risk parameters (PD and LGD) as well as in the underlying network structure (distribution of exposures)[41]. We take Q4-2019 as baseline year (pre-pandemic period) and we let first LGD parameters by sector and country vary over time according to the actual estimates, and we keep unchanged relative to Q4-2019 the PD parameters and the network structure. In this way, by subtracting the new derived losses (mean and 99th percentile) with those obtained for Q4-2019, we derive the contribution of LGD parameters from Q1-2020 onwards relative to Q4-2019. We repeat the exercise other two times by letting vary also the PD parameters to compute the PD contribution, and then the network structure, keeping always unchanged the other two risk dimensions. Finally by summing each factor loss contribution to the actual loss in Q4-2019 and by taking the difference of it with the

[41] We keep the correlation structure unchanged and benchmarked with the baseline case for consistency and comparability purposes.

actual loss experienced in each specific quarter (given all factors changing simultaneously) gives us the interaction term contribution, which is a function of the combined variations in LGDs, PDs and the network structure.

Figure 13 provides insights on the contribution to expected losses of each risk factors (PD, LGD, and NTW) and the interaction term. First of all, we can notice that the relative contribution change over time. In Q1-2020 relative to Q4-2020, the main contributor to the increase in expected losses is the network structure, explain 87% of the total loss variation in between Q4-2019 and Q1-2020. Though, we can see that in Q2-2020, at the peak of crisis, counterparty default risk captured by the variations in PD parameters became the most important factor with 65% of total contribution (relative to Q4-2019), followed by the network structure with a positive contribution of 25%, and by the interaction term (10%). We can see that this trend persists throughout the whole pandemic period till the last quarter available Q4-2021, with an increasing PD loss contribution. This finding emphasizes the role played by counterparty default rates in the build-up of average risk during the pandemic relative to the pre-pandemic period. Interestingly we can notice that LGD loss contribution turned negative, emphasizing that not all the risk factors contributed to a deterioration in the expected loss estimate.

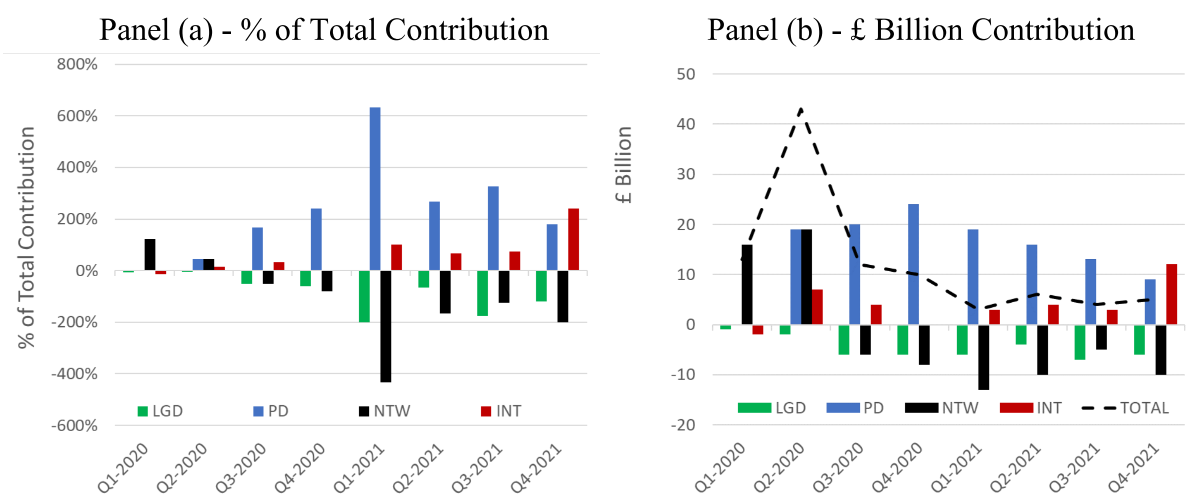
Figure 13: Risk factors Contribution to Expected Losses



Note: "LGD" refers to the contribution made by quarterly changes in Loss-given-default parameters relative to Q4-2019, "PD" refers to the contribution made by quarterly changes in counterparty probability of defaults, "NTW" refers to the contribution made by quarterly evolution of the network structure, and "INT" to the contribution of the interaction term, which captures the interaction effects of the simultaneous change of LGD, PD, and NTW parameters.

Moving to the analysis of tail risk, approximated by the 99th percentile of the loss distribution, we can notice that the contribution of each risk factor to changes in tail losses varies remarkably across quarters. Although initially the network structure has a positive contribution alike for expected losses, its contribution becomes negative reaching a peak in Q1-2021, and remains negative till Q4-2021. This implies that the conformation of the network structure is risk absorber relative to the conformation in Q4-2019. Contrary, counterparty default risk (PD) is the major driver to the overall increase in tail risk, compensating the negative contribution of both variations in the network structure and in LGD parameters. Moreover, the interaction term becomes also positive and relevant especially in Q4-2021 (major contributor), emphasizing that the current risk environment (NTW, PD, LGD) is still less beneficial than the one experienced in Q4-2019. Interestingly we can notice that the network structure contribution is positive for measuring expected losses, while negative for tail risk. Overall, counterparty default risk remains the key contributor to higher expected and tail losses relative to the pre-pandemic period.

Figure 14: Risk Factors Contribution to Tail Losses (99th percentile)



Note: "LGD" refers to the contribution made by quarterly changes in Loss-given-default parameters relative to Q4-2019, "PD" refers to the contribution made by quarterly changes in counterparty probability of defaults, "NTW" refers to the contribution made by quarterly evolution of the network structure, and "INT" to the contribution of the interaction term, which captures the interaction effects of the simultaneous change of LGD, PD, and NTW parameters.

5.6 Limitations

We want to emphasize that our network approach to stress testing the UK banking system' asset side have some limitations, which in turn have implications for our CaR and CCaR estimates, with potential revisions both upwards and downwards. First of all, we only model losses from direct exposures, and we don't model amplification effects or second-round losses arising from interbank or intra financial sector's exposures via contagion or fire-sales spillovers. In this respect, our approach is closer to a standard stress test methodology rather than a financial network stress testing methodology as developed by Montagna et al. (2021), Sydow et al (2021), and Roncoroni et al. (2021), among others. As we have seen CaR and

CCaR estimates capture a little amount of losses stemming from credit institutions, accounting for 1% of losses in extreme stress events. Hence, our estimates tend to underestimate systemic risk spillovers stemming from this sector. Future extensions of this work should tackle this methodological gap, and directly model financial amplification mechanisms in the interbank market. This is even more important since losses stemming from security exposures are modelled similarly to credit risk exposures, and are not derived with an ad-hoc price impact function as for instance in Coen et al. (2019). Our estimates are conservative in this respect, since fire-sales mechanisms via overlapping portfolios of exposures may also negatively affect the price of securities of those counterparties which actually did not default in tail scenarios. Hence, banks' conditional capital at risk estimates should be a function of correlations in the real economic sector as well as in the financial sector since the price of securities is actually a function of financial institutions' management strategies. In this respect, modelling financial amplification mechanisms would benefit from extending the coverage of the Global Network beyond the banking sector, for instance to the UK insurance and pension fund sectors as well as to the investment fund sector. We leave these dimensions to future research.

Another limitation of the current work is that it relies on PD parameters that are not counterparty-specific. Our risk parameters are homogeneous across counterparties belonging to the same sector and country. This assumption which is due to the limitation in terms of granular coverage of our supervisory data source may lead to an over estimation of CaR and CCaR estimates. The rationale is the following. Given that the distribution of exposures follow a power-law distribution, that is, large corporates capture 90% of total UK banks' exposures, and given that small corporates are likely to experience on average higher probability of defaults than large corporates, we tend to assign higher PDs to the set of largest exposures. In this respect, the current work would benefit from using counterparty-specific PD parameters at least for a set of large quoted corporates so as to increase the degree of heterogeneity in counterparty risk.

Related to this, we have to emphasize that our data source for risk factors exploits banks' own assessment of their specific counterparty risk. This subjective assessment may present both advantages and disadvantages. On the one hand, banks may have superior knowledge on their counterparty risk compared to external data sources, thereby contributing to enhance our loss estimates. Moreover, banks' counterparty risk factors reflect banks' short to medium term expectations on financial and economic developments (1-year ahead). Hence, our CaR and CCaR estimates incorporate banks' expectations about the future state of the world economy, thereby working as potential early warning indicators for a banking system's loss of confidence (or overconfidence). This type of signals may inform policy makers on the timing of releasing (tightening) capital buffer requirements in order to sustain funding to the corporate sector. This is crucial since a deterioration of confidence vis-à-vis the non-financial corporate sector may directly exacerbate liquidity conditions of non-financial corporates, and in turn increase the likelihood of corporate defaults, which in the end worsen banks' solvency position. On the other hand, using subjective estimates of risk factors may lead to potential bias, since banks may under or overestimate counterparty risk. In this respect, we tackle this source of estimation bias by averaging PD and LGD parameters by country and sector across reporting banks, thereby implementing a pool approach to counterparty risk. Nevertheless, for policy analysis and policy applications it is important to test our results to homogenous and heterogeneous variations in risk parameters so as to overcome parameter uncertainty, as we did for scenario uncertainty. However, CaR estimates are a linear function of both PD and LGD parameters, hence a homogeneous increase in risk factors would lead to a homogenous increase in expected losses. This is not true for CCaR estimates since an increase in PD parameters would affect also the correlation structure of corporate defaults. In this respect, we tested our results to a variation in the correlation structure, risk parameters, and network structure.

In the end, we want to highlight few other concerns and potential extensions of our work. Our results are not affected by the number of Montecarlo simulations. We perform sensitivity analysis using 10.000/30.000 simulations, and CCaR estimates do not change. Finally, the granular coverage of the Global Network, although quite comprehensive, still requires further improvement. Roughly 57% of total exposure amounts were captured with aggregated exposures by country and sector, thereby not entering into the calculation of the stochastic loss component for the Conditional Capital at Risk measure. Instead they enter the calculation of the CCaR as expected loss component, thereby leading to an under-estimation of the severity and likelihood of tail events.

6. Conclusion

Concentration risk and interconnectedness are systemic risk dimensions that can only be unravelled through network analysis. Shocks to large corporates, the granular origins of aggregate fluctuations (Gabaix 2011), and a high level of interdependency in the intersectoral input-output linkages, the network origins (Acemoglu et al. 2012), may drive aggregate fluctuations in output. The same rationale applies to the microstructure of the financial system, which in turn determines fluctuations in Systemic Risk (Acemoglu et al. 2015). The degree of concentration risk as well as of interconnectedness both in the economic system and in the financial system have strong implications for the system's functioning and its propensity to (in)stability. In this regard, we provide evidence that the microstructure of the UK banking sector is potentially conducive to instability given its highly concentrated and interconnected network structure.

Hence, in this set-up, idiosyncratic shocks to specific UK banks' counterparties have implications for fluctuations in the aggregate level of systemic risk, and likewise idiosyncratic shocks to specific banks may have implications for fluctuations in economic activity. Our work focused on the former, by assessing systemic risk from a system-perspective as a function of the actual network structure of UK banks' exposures and of variations in counterparty risk conditions proxied by changes in PD and LGD parameters of banks' counterparties. Hence, we computed UK banks' expected losses also defined as Capital at Risk (CaR) which amounted to £31 billion in Q4-2021, lower compared to Q2-2020 - the peak of the Covid-19 crisis - (£37 billion), though still higher relative to the pre-crisis period (£27 billion).

However, CaR estimates are not a function of the network structure, and do not take into consideration either the distributional features of the network (concentration risk and interconnectedness) or the inter-sectoral input-output linkages among banks' counterparties. For this reason, we modelled scenario uncertainty as a function of the realized network structure. By means of Monte Carlo simulations we thus compute the Conditional Capital at Risk measure (CCaR), which quantifies the severity and likelihood of tail events. We estimate that at the peak of the Covid-19 pandemic the probability of experiencing an extreme stress event in the UK banking sector of more than £91 billion losses was close to 4.1%, up from 1% in the pre-crisis period.

Furthermore, we shed light on the sensitivity of our results to a strengthening in the degree of correlation of counterparty defaults, that is, a tightening in the intersectoral input-out linkages. Conditional to this stressed correlation structure, the probability of experiencing extreme stress events in Q2-2020 further increase to 6%. The intuition behind this result is that the strengthening of the correlation in counterparty defaults leads to a higher likelihood of experiencing clusters of defaults by country and sector thereby increasing the severity of tail events and so their conditional likelihood. The CCaR decomposition by sector and geographical region highlights that the UK banking sector is more exposed to systemic risk spillovers stemming from exposures outside the UK jurisdiction as to within the UK jurisdiction, and this trend is especially visible in recent quarters.

Overall, we provide evidence that tail-risk (CCaR) in the UK banking sector increased remarkably more than the average risk (CaR), thereby implying that expected losses are not a good proxy for tail losses or systemic risk. Thus, a stochastic granular stress testing approach is needed to take into account uncertainty over the realization of the macro stress scenario. Moreover, also the sources of expected losses differ from those of tail losses, with the network structure giving a positive contribution to average losses, while acting as shock absorber in the tail.

To conclude, we highlighted at the beginning of this paper the challenges the research and policy community have faced in their quest to shed light on the role that modern financial systems play in the unfolding of events like the Great Financial Crisis. By adopting a network perspective, empirical work has started to investigate the microstructure of the financial system across countries and sectors focusing on those features - concentration and interconnectedness - that may contribute to the financial system's functioning and its propensity to (in)stability. In this paper, we have highlighted the role those features play in strengthening the severity and likelihood of tail events for the UK banking sector. In particular, we determine that highly interconnected and concentrated financial and economic systems may provide clear efficiency benefits during tranquil times, but at the same time the very same features also increase the system's level of instability once the reign of tranquillity ends.

References

Abuzayeda B., Bourib El., Al-Fayoumia N., Jalkhc N. (2021), "Systemic risk spillover across global and country stock markets during the COVID-19 pandemic", *Economic Analysis and Policy* 71:180-197

Acemoglu D., V.M. Carvalho, A. Ozdaglar, and A.Tahbaz-Salehi (2012), "The Network Origins of Aggregate Fluctuations", *Econometrica* 80(5): 1977-2016

Acemoglu D., A. Ozdaglar, and A.Tahbaz-Salehi (2015), "Systemic Risk and Stability in Financial Networks", *American Economic Review* 105(2): 564–608

Adrian T., N. Boyarchenko, and D. Giannone (2019), "Vulnerable Growth", *American Economic Review* 109(4): 1263-1289

Adrian T, Brunnermeier MK (2016), "Covar", *American Economic Review* 106(7):1705–41

Aikman, D., Chichkanov, P., Douglas, G., Georgiev, Y., Howat, J. & King, B. (2019). Systemwide stress simulation. Bank of England Staff Working Paper No. 809.

Aldosoro and Alves (2018), "Multiplex interbank networks and systemic importance: An application to European data" *Journal of Financial Stability* 35:17-37

Alessi, L. and Battiston, S. (2021), "Two sides of the same coin: Green Taxonomy alignment versus transition risk in financial portfolios", Working Paper, European Commission – Joint Research Centre, forthcoming.

Alessi, L., Battiston, S., Melo, A.S. and Roncoroni, A. (2019), "The EU Sustainability Taxonomy: a Financial Impact Assessment", Technical Report, European Commission – Joint Research Centre.

Alessi L., Battiston S., and Melo S. (2021), "Travelling down the green brick road: a status quo assessment of the EU taxonomy", ECB Macroprudential Bullettin 15-2021

Allen F., Covi C., Gu X., Kowalewski O. and Montagna M. (2020) "The Interbank Market Puzzle" Bank of England Staff Working Paper No. 862

Allen F., and A. Babus. (2008), "Networks in finance", Wharton Financial Institutions Center Working Paper No. 08-07

Allen F. and D. Gale (2001), "Comparative Financial Systems: A Survey", Wharton School Center for Financial Institutions, University of Pennsylvania.

Anand K., B. Craig and G. von Peter (2014), "Filling in the Blanks: Network Structure and Interbank Contagion", Discussion Paper No. 2, Deutsche Bundesbank

Bardoscia M., P. Barucca and A.B. Codd, and J. Hill (2019a) "Forward-looking solvency contagion" *Journal of Economic Dynamics and Control*, 108:1-26

Bardoscia M., G. Bianconi, G. Ferrara (2019b), "Multiplex network analysis of the UK over-the-counter derivatives market", *International Journal of Finance and Economics* 24(4): 1520-1544

Bardoscia M., P. Barucca, S. Battiston, F. Caccioli, G. Cimini, D. Garlaschelli, F. Saracco, T. Squartini, and G. Carldarelli (2021), "The Physics of Financial Network", *Nature Reviews Physics*, 3: 490-507.

BCBS (2014), "Supervisory framework for measuring and controlling large exposures". Standards, Basel Committee on Banking Supervision

BIS (2021) "International Banking and Financial Market Developments", BIS Quarterly Review, March 2021.

BIS (2013), “Global Systemically Important Banks: Updated Assessment Methodology and the Higher Loss Absorbency Requirement” Basel Committee on Banking Supervision

Brownlees C. and R. Engle (2017), “SRISK: a Conditional Capital Shortfall Measure of Systemic Risk”, ESRB Staff Working Paper No. 37

Coen J., C. Lepore, and E. Schaanning (2019) “Taking regulation seriously: fire sales under solvency and liquidity constraints”, Bank of England Staff Working Paper No. 793

Covi G., and MZ Gorpe, C. Kok (2021), “CoMap: Mapping Contagion in the Euro Area Banking Sector”, *Journal of Financial Stability* 53

Craig B. and Goetz von Peter (2014) “Interbank Tiering and Money Center Banks”, *Journal of Financial Intermediation* 23:322-347

Degryse H. and G. Nguyen (2007) “Interbank Exposures: An Empirical Examination of Contagion Risk in the Belgian Banking System”, *International Journal of Central Banking* 3 (2): 123-71.

Düllmann, Klaus, Martin Scheicher, and Christian Schmieder, 2008, “Asset correlations and credit portfolio risk - an empirical analysis,” *Journal of Credit Risk* 4(2), pp. 37–62.

ECB (2019), “ECB guide to internal models”, October 2019:

EIOPA (2018), “Financial Stability Report”. December.

Elsinger H., A. Lehar, and M. Summer (2006), “Risk Assessment for Banking Systems”, *Management Science* 52 (9): 1301–14

Eurostat (2008) “NACE Rev. 2 Statistical classification of economic activities in the European Community”, Accessed at: https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF

François L., Kaminska B., Aparicio C., Dworak B., Garcia L., Durant D., Ristori L., Kirchner R., Calleja R., Katsigianni E., Vitellas D., Pistelli F. (2021), “The benefits of the Legal Entity Identifier for monitoring systemic risk”, ESRB Occasional Working Paper No. 18

Financial Stability Board (2019), “Thematic Review on Implementation of the Legal Entity Identifier”, Peer Review Report, 29th March 2019

Gabaix X. (2011), “The Granular Origins of Aggregate Fluctuations”, *Econometrica* 79(3):733-772

Gai P., Haldane A. and S. Kapadia (2011). Complexity, concentration and contagion. *Journal of Monetary Economics* 58, 453–470.

Geithner, T. (2014), Stress Test: Reflections on Financial Crises, Crown Publishing Group, Random House

Glasserman, P. & Young, H. P. (2016). Contagion in financial networks. *Journal of Economic Literature* 54, 779–831.

Glasserman P. and H.P. Young (2015) “How likely is contagion in financial networks?”, *Journal of Banking and Finance*, 50:383-399

Halaj G. and C. Kok (2013), “Assessing Interbank Contagion Using Simulated Networks”, *Computational Management Science* 10 (2):157-86

Hardy D. and C. Schmieder (2013): “Rules of thumb for bank solvency stress testing”, IMF Staff Working Paper No 13/232

Huser A-C., C. Lepore, and L. Veraart (2021) “How does the repo market behave under stress? Evidence from the Covid-19 crisis” Bank of England Staff Working paper No. 910

Huser A-C. (2015) “Too interconnected to fail: a survey of the interbank networks literature”, *Journal of Network Theory and Finance* 1(3):1-50

Laurent F., B. Kaminska, C. Aparicio, B. Dworak, L. Garcia, D. Durant, L. Ristori, R. Kirchner, R. Calleja, E. Katsigianni, D. Vitellas, F. Pistelli (2021) “The benefits of the Legal Entity Identifier for monitoring systemic risk” ESRB Occasional Paper No. 18

Lopez J. A., (2004), “The empirical relationship between average asset correlation, firm probability of default, and asset size,” *Journal of Financial Intermediation* 13(2): 265–283.

Montagna M., Torri G., Covi G. (2021) “On the Origin of Systemic Risk”, Bank of England Staff Working Paper No. 906

Newman M. (2004), “Power laws, Pareto distributions and Zipf’s law, *Contemporary Physics* 46(5): 323-351

Rees D. and Rungcharoenkitkul P. (2021), “Bottlenecks: causes and macroeconomic Implications”, BIS Bulletin No. 48

Roukny T., S. Battiston and J.E. Stiglitz (2018) “Interconnectedness as a source of uncertainty in systemic risk”, *Journal of Financial Stability* 35:93-106

Schrimpf, A., Shin, H. S. & Sushko, V. (2020). Leverage and margin spirals in fixed income markets during the COVID-19 crisis. Bank for International Settlements, BIS Bulletin No 2.

Sheldon G., and M. Maurer (1998), “Interbank Lending and Systemic Risk: An Empirical Analysis for Switzerland”, *Swiss Journal of Economics and Statistics* 134 (4): 685-704

Sydow M., Schilte A., Covi G., Deipenbrock M., Del Vecchio L., Fiedor P., Fukker G., Gehrend M., Gourdel R., Grassi A., Hilberg B., Kaijser M., Kaoudis G., Mingarelli L., Montagna M., Piquard T., Salakhova D., Tente N. (2021), “Shock amplification in aninterconnected financial system of banks and investment funds”, ECB Staff Working Paper No. 2581

Taleb N., E. Canetti, T. Kinda, E. Loukoianova, and C. Schmieder (2012), “A New Heuristic Measure of Fragility and Tail Risks: Application to Stress Testing,” IMF Working Paper no. 12/216.

Upper C., 2011, “Simulation Methods to Assess the Danger of Contagion in Interbank Markets”, *Journal of Financial Stability* 7 (3): 111-25

Van Lelyveld I. and F. Liedorp, (2006) “Interbank Contagion in the Dutch Banking Sector: A Sensitivity Analysis”, *International Journal of Central Banking* 2 (2): 99-133.

Appendix A - Data

1. A Roadmap to Construct a Granular Data Infrastructure

The Basel Committee on Banking Supervision (BCBS, 2014) set out a global regulatory framework called “large exposures” to monitor and measure concentration risk arising from a counterparty failure. This was a necessary step in raising global awareness among financial market participants (public and private) about the need for more granular and up-to-date sources of exposure information. This need is now even more compelling in light of the sudden and fast-moving Covid-19 Pandemic that has affected counterparties across sectors and countries in remarkable different ways and with different timings.

A successful implementation of a comprehensive data collection on exposures at the firm-counterparty level has been hard to achieve, and is yet to be fully accomplished. While the regulation in theory requires firms to precisely identify counterparties via their name and legal entity identifier (LEI code), in practice the consistency of the way in which counterparties were identified could vary significantly across regulated firms. This counterparty identification problem via LEI code and firms’ names is not exclusive for the UK jurisdiction, but is shared across jurisdictions as highlighted for the Euro Area (EA) in Covi et al. (2021). Counterparty legal names differ across reporting institutions even within the same jurisdiction, and LEI codes are often difficult to retrieve and associate to counterparty names. In general, the LEI adoption across jurisdictions has been unbalanced and is low outside securities and OTC derivatives markets (FSB, 2019: 25), making it difficult for credit institutions and financial firms to associate LEI codes to counterparties especially from the banking book. Hence, LEI coverage still remains too low to encourage further regulatory uses or to reach a potential tipping point where voluntary take-up by market participants would suffice to propel further adoption (FSB, 2019: 25).

These data challenges haven’t held back the implementation of the large exposures regulation or the collection of other granular exposure datasets, such as security or derivative portfolio holdings, though they have slowed down and complicated the exploitation of this exposure-type datasets for monitoring financial stability risks and for supervisory purposes. For central bankers, supervisors, and regulators to maximise the potential of using these datasets, the first and foremost pre-requisite is a high-quality mapping of counterparties to legal entity names and LEI codes in a consistent manner across firms and time, which is not yet the case in current granular regulatory data collections.

In 2012, the Financial Stability Board - conscious of the challenges and costs for the private sector as well as the wide range of benefits[42] this data transition and improvement in underlying data quality may provide in the medium-long run to all market participants - was endorsed by G-20 countries to promote a "global adoption of the LEI framework to support authorities and market participants in identifying and managing financial risks". As part of this initiative, the FSB in 2014 kick-started the construction of the Global Legal Entity Identifier Foundation (GLEIF) database as the operational arm of the system that federates local LEI issuers under the oversight of the LEI Regulatory Oversight Committee (ROC)[43]. This presents a fundamental step in disseminating reference and group structure information as a public good and in pushing forward and supporting a global-scale adoption of LEI codes for the identification of financial and non-financial firms. As emphasized by the ESRB (Laurent et al. 2021), LEI codes have the potential to become the identifier of the global economy. Moreover, in order to support LEI adoption by financial market participants, the EU Regulation No 600/2014 (also defined as MiFIR) has limited the access entities can have to financial markets in the European Union without an associated LEI code. As documented by the ESRB, the adoption in Q1-2017 of this LEI reference standard which was applied to various collections of granular data sources[44] led to a remarkable increase in LEI issuance per month, that is, in counterparty LEI-based identification. Overall these regulatory and data initiatives represent a key milestone in terms of supporting a global scale adoption of LEI codes as firm-specific identifiers.

Nevertheless, a common and consistent adoption of LEI codes across firms on a global scale still requires a full integration of the LEI identification approach into data management systems, both in the private sector as well as in the public sector. Why is this integration process not straightforward? The current data management systems are based upon using unique internal identifiers (*not* LEI-based) to classify each counterparty (client), and those internal IDs differ across reporting banks, even within the same jurisdiction. In order to bridge this data gap, the GLEIF database represents an excellent master data reference source for mapping

[42] In terms of i) monitoring financial risks; ii) exposure aggregation in data reporting; iii) statistical analysis; iv) understanding the structures of multinational companies, market structure and trading networks; v) and facilitating market surveillance and compliance assessments (FSB, 2019: 25).

[43] The ROC is composed by 65 financial markets regulators and other public authorities and 19 observers from more than 50 countries. It promotes the broad public interest by improving the quality of data used in financial data reporting, improving the ability to monitor financial risk, and lowering regulatory reporting costs through the harmonization of these standards across jurisdictions. Access at: https://www.leiroc.org/

[44] The European Market Infrastructure Regulation (EMIR), the Markets in Financial Instruments Directive (MiFID II), MiFIR, Solvency II and the Central Securities Depositories Regulation (CSDR). See Laurent et al. (2021) for an overview of these regulations and impact.

counterparty entities with LEI codes by exploiting the legal name and/or the security identifier (ISIN code) stored in the GLEIF database[45].

Capitalising on this master data reference source, we provide a data infrastructure framework for cleaning, mapping, and merging firms' counterparty information and for reconstructing balance sheet based granular financial network exploiting multiple supervisory data sources. As part of this framework, we apply a mapping-searching algorithm across 8 million entities collected from the entire spectrum of global counterparties in the UK banks' portfolios of loan and security exposures. We thus document a roadmap to build a granular automatic and scalable data infrastructure. This is a necessary step for then studying and analysing the network of UK banks' exposures on a global scale exploiting counterparty-specific information. In this respect, we follow the works of Covi et al. (2021) in mapping bank-specific information using LEI identifiers and Montagna et al. (2021) who extend the approach to a global map of counterparties in order to merge different granular supervisory datasets for the Euro Area (EA) banking sector. The datasets we use share the standard structure of granular datasets collected both for supervisory and non-supervisory purposes. They also share the standard structure for both security ISIN-based datasets as well as for LEI-based datasets, making this data infrastructure framework applicable to any data management system. The data infrastructure's value-add is therefore twofold, respectively enabling i) the mapping of counterparties via LEI and ISIN codes and so producing a consistently mapped network of exposures, and ii) the linking of qualitative and quantitative counterparty-specific information via LEI codes using multiple private and public data sources. Overall, the data infrastructure and mapping algorithm increase the coverage of counterparty LEI codes from 55% to 91% for the large exposures dataset (LE) and up to 97% and 87% for the security ISIN-based datasets from respectively 0% and 66%.

1.1 Network Data Sources

In this section we describe the datasets we use to construct the UK banks' global network. In order to achieve a comprehensive coverage of the UK banks' asset side we need to draw from various data sources of different levels of granularity. First, we introduce two *entity-based datasets* (LE and C67) capturing entity to entity relationships, where the reporting and counterparty sides are identified via Legal Entity Identifiers (LEI codes). Next, we present two *security-based* datasets covering entity to security relationships, where the reporting side is still

[45] As explained before, for security and derivative instruments, a full map of ISIN codes to LEI codes is provided by the GLEIF website. So for this type of assets, the LEI mapping process is easier to implement.

identified via a LEI code, as in the entity-based datasets, but the counterparty side is now identified via ISIN codes tracking the security issued. The level of granularity of the security-based datasets is higher than the entity-based datasets since multiple ISIN codes belong to one single entity. In the end, since the coverage of these granular datasets is not exhaustive, we collect one *aggregate-based* dataset, for which the reporting side is identified with a LEI code, whereas the counterparty side is aggregated by sector and country of origin.

1.1.1 Granular Entity-Based Datasets

Large Exposures Dataset (LE)

The large exposures (LE) framework serves as a backstop to the capital framework by ensuring that the maximum loss a bank would face in the event of a sudden default of a counterparty or a group of connected clients that are linked by economic dependence or control would not endanger the bank's solvency.[46] All exposures captured under the risk-based capital framework are subject to the LE framework, including off-balance sheet exposures. This includes exposures in both the banking and trading book without applying risk weights or degrees of risk. Banks are required to limit their exposures to an individual client or group of connected clients to 25% of Tier 1 capital, although some exposures might be subject to different limits or exemptions[47]. Banks may use eligible credit risk mitigation techniques to reduce these exposures.

The supervisory data collection on exposures in scope of the LE framework started in 2014 as part of the Common Reporting Framework (COREP) regime. Specifically, COREP templates C.27, C.28, C.29, C.30 are submitted by UK banks on a quarterly basis. Banks are required to submit these returns on an individual entity basis, on a consolidated and, if applicable, sub-consolidated basis (as is the case for ring-fenced banks for example). The LE dataset captures only those exposures that meet the definition of a large exposure i.e. that are larger than 10% of a bank's/consolidation group's Tier 1 capital or on a consolidated basis, are above £260 million[48]. The LE dataset is constructed with an entity to entity relationship. Each counterparty is identified with the legal name or group name and where possible with a Legal Entity Identifier (LEI code). Additional counterparty information is provided such as the

[46] Supervisory framework for measuring and controlling large exposures (bis.org)

[47] From 2014 to 2017, transitional provisions outlined in CRR Article 494 allowed banks to include decreasing amounts of Tier 2 capital in the definition of eligible capital as it applied to large exposures rules. In 2014, the applicable amount of Tier 2 capital was capped at a 100% of Tier 1 capital. During 2015 and 2016, the applicable amount of Tier 2 capital was capped at 75% and 50% of the value of Tier 1 capital respectively. From 1 January 2017, the applicable amount of Tier 2 capital was capped at a third of Tier 1 capital as set out in CRR Art 4 (71)(b). From 1 Jan 2022, eligible capital is defined as Tier 1 capital only as per the Basel Standard on large exposures.

[48] Prior to 1 Jan 2022, this threshold was set at €300 million.

SECTOR, COUNTRY (of incorporation) and NACE classification of the entity. The LE dataset also provides a rich set of exposure attributes, which allow us to distinguish debt, equity, derivate and off-balance sheet exposures. Moreover, the dataset also provides information on the amount of exposure which is subject to exemptions from LE limits and the amount of exposure that is secured by credit risk mitigation instruments[49]. For our scope, we will focus on two main variables of interests: gross original exposure amount and the gross original exposures amount after having deducted the credit risk mitigation instruments so as to classify the exposure amounts into secured and unsecured exposure amounts. Prior to January 2022, the LE dataset also provided information on the maturity breakdown of the top 10 largest exposures vis-à-vis regulated and unregulated entities (template C.30).

The LE dataset, given its reporting threshold, is quite comprehensive covering banks' exposures towards credit institutions (CI), governments (GG) and central banks (CB) and to a lesser extent, exposures towards non-financial corporations (NFC) and non-bank financial corporations (FC). This difference in the sectoral coverage is due to the average size of the entities within those sectors. It is expected that, on average, NFCs and FCs would consist of small-medium size entities and therefore, a larger share of exposures to these entities are expected to be below the reporting threshold. Looking at the interbank network specifically, exposures from large reporting banks towards small banks also tend to be under-populated. No exposures in relation to the the household sector (HH) are provided as part the LE dataset. The dataset is a global dataset, capturing UK banks' large exposures vis-à-vis entities worldwide. The dataset is UK-centric, meaning that on the reporting side, only UK firms, subsidiaries of international banking groups domiciled in the UK or UK consolidation groups are present. This feature is common across all data sources.

Although the LE dataset is very rich, constructing a consistent network of bilateral relationships is not immediate. Consistent identification of counterparty entities is complex and challenging. Reporting banks may classify the very same counterparty with a slightly different name, or report the legal name of the subsidiaries instead of the group name. Moreover, the LEI dimension is not always available because banks may not have a LEI code (especially for small firms or subsidiaries) or because the reporting bank does not provide the LEI code against the counterparty (indicated as "not available/NA"). Moreover, when the entity on the counterparty side is classified as a group of connected clients, the reporting bank is obliged to

[49] For example, exemptions are applied to exposures vis-à-vis certain general governments and central banks based on their assigned risk weights as per relevant regulatory requirements as well as to certain entities in bank's wider group based on prior approval from the PRA.

report only the name of the head of group but not the related LEI code. There may also be variation in LEI codes and legal names reported against the same counterparty by the same reporting bank across time. This identification gap on the counterparty side is the main challenge we need to overcome in order to construct a consistent network of bilateral relationships. Similar problems arise for the COUNTRY, SECTOR and NACE dimensions which are often missing. These data identification issues are common to all granular data sources since the collection of these datasets started only relatively recently and reporting infrastructure is still being refined. In this respect, one of the main contributions of this paper is to develop and present a procedure to overcome these challenges and fully exploit these extremely rich and unique data sources. This data source covers over time roughly as unbalanced panel 425 reporting banks, 6592 counterparties for a total of 213.428 exposure data up to Q4-2021.

Liability Dataset (C.67)

The Liability Dataset (C67) provides information on the 10 largest funding sources of the UK banks on a monthly frequency where the funding obtained from each counterparty or group of connected clients exceeds a threshold of 1% of total liabilities as at the reporting date. Its collection aims at monitoring concentration risk on the liability side and allows regulatory authorities to monitor a bank's liquidity risk that falls outside the scope of the reports on liquidity coverage and stable funding. This dataset is useful to reconstruct a part of the asset side of non-UK bank exposures, and to complement to certain extent exposures from large-size UK banks towards small-size UK banks which are not captured in the LE dataset due to the reporting threshold. The dataset also captures exposures from governments and central banks towards UK banks.

The dataset presents a similar structure, although simplified, to the LE dataset. Exposure amounts are reported in gross terms, and exposures are classified by type of instruments, which allows us to disentangle between secured and unsecured funding exposures[50]. A maturity breakdown is also provided, with a variable defining the average weighted maturity of the exposure in days. The funding provider is also identified with their legal name and the related LEI code. The COUNTRY and SECTOR dimensions are also reported, but not the NACE classification. This dataset presents the same data identification issues previously described thereby also requiring a consistent mapping of the counterparty entities. In this respect, this

[50] Information on the Large Exposures data can be retrieved from the Prudential Regulation Authority's Website.

data source covers over time roughly 241 reporting banks, 2201 counterparties and 84506 data points up to Q4-2021.

1.1.2 Granular Security-Based Datasets

In order to monitor and model market risk and increase the coverage of the UK banks' Global Network, we aim to collect UK banks' security holdings. To achieve that, we exploit two main and complementary data sources, respectively the SHS and AS datasets. The former is collected on an annual basis by the BoE's Stress Test Division to exclusively monitor seven major UK banks (ACS banks) on a consolidated basis for the annual cyclical scenario[51], whereas the latter is collected on a quarterly frequency by the BoE's Data and Statistics Division to monitor a larger sample of UK banks on an unconsolidated basis. Both datasets have a common data structure. Reporting banks are identified with their legal names and LEI codes, whereas the counterparty side is mapped with an ISIN code for each security held, an LEI code of the issuer, and a legal name as well as the country and sector of the counterparty. For each security, the datasets also report the maturity date of the contract, splitting between equity and debt securities, and the currency in which the security has been issued. In the end, for each security exposure, amounts are reported at nominal and market values. These datasets are big granular datasets. The SHS and AS datasets covers over time respectively 13 and 33 reporting banks; 28,033 and 41,197 counterparties; 69,938 and 115732 ISIN-based security instruments; and 199,501 and 1,360,809 security exposures up to Q4-2021.

Just as with the entity-based datasets presented in the previous sub-section, these security datasets also present some data issues. For instance, ISIN codes might be wrongly typed and LEI codes identifying the issuer of the security may be missing as well as the country and sectoral tags. Moreover, counterparty names are also reported differently across reporting firms. Hence, in order to construct a multilayer network of granular exposures in which reporting and counterparty entities are consistently and uniquely identified across reporting firms and time, these data issues need to be tackled and solved via a cleaning a mapping procedure which will be presented in section 1.2.

1.1.3 Aggregate Sector-Country Exposures

With the LE, C67, SHS and AS datasets we are able to map entity to entity relationships. They represent the granular information of our global network. However, the coverage of UK banks' total assets is not complete if we limit our analysis to this collection. For the objective of

[51] ACS banks are those banks that are subject to the annual cyclical scenario, and there are seven of them: HSBC, Barclays, Standard Chartered, Lloyds, Nationwide, Santander, and Royal Bank of Scotland.

performing a stress test exercise and computing accurate and reliable financial stability metrics from a supervisory and policy perspective, we should aim to increase the coverage of the asset side to the largest extent possible even though we may lose some degree of granularity in doing so. In this respect, we collect data from FINREP supervisory template F.20.04, which contains information on the geographical and sectoral breakdown of UK banks' assets. That is, for each country in which the reporting bank is exposed and for each asset class (derivatives, equity instruments, debt securities, and loans and advances) and for each sector as applicable (credit institutions - CI, other financial corporations - FC, non-financial corporations - NFC, central banks - CB, general governments - GG, and households - HH), the firm is required to report its exposures as gross carrying amounts. These three dimensions, COUNTRY, SECTOR and type of INSTRUMENT are consistent with the dimensions previously described in the granular template thereby not requiring us an additional matching effort across datasets. Our wrangled F.20.04 dataset, also defined as aggregate exposure dataset, has 166,436 observations on 11 attributes, starting in Q1 2018 and (for our purposes here) ending in Q3 2021.

This dataset is important for the reconstruction of the Global Network especially for filling up three main counterparty sectors: HH, NFC, and FC. On the one hand, the household sector is missing from the granular datasets previously reported. This sector is crucial for modelling risk stemming from domestic exposures since a large share of exposures to households are within the UK. On the other hand, UK banks' large exposures vis-à-vis non-financial corporates and non-bank financial corporates may miss a significant portion of exposures towards small-medium enterprises and funds due to the reporting threshold of the large exposure regulation. By contrast, large exposures towards CIs, GGs and CBs tend to be sizeable enough to be captured in the LE dataset. Therefore, the aggregate exposure dataset helps to extend the coverage of these remaining sectors. Moreover, this dataset is also useful for benchmarking purposes so as to quantify the share of granular exposures captured by country and sector.

After having collected the F.20.04 dataset, we need to proceed with the data manipulation procedure since the granular datasets and the aggregate exposure dataset cannot be directly merged as they stand. Both the exposure-datasets and the aggregate-exposure dataset need to be first cleaned, mapped and subset before they are merged since these datasets have exposures in common. This is an important step to avoid double-counting issues. In the next section, we are going to describe the procedure for cleaning and mapping the granular exposure-datasets, whereas in section 1.4 we provide a description of the merging procedure.

1.2 Cleaning and Mapping Algorithm

1.2.1 Cleaning Procedure

In order to construct the UK banks' network of granular exposures, we need to identify reporting banks and counterparty entities uniquely across three main dimensions: time, dataset, and reporting firm. In the LE and C67 datasets we have two main identifiers, respectively the NAME and LEI of the counterparty. However, the very same counterparty may be reported with a slightly different name even by the same reporting bank across datasets and time. The name of the counterparty may also differ across reporting banks within the same dataset and reporting period. In order to homogenise the name, we create an additional variable, defined as CLEAN_NAME, which is the original name cleaned of special characters, white spaces, and digits. Next, we move to the LEI identifier, which is a unique code of 20 characters. However, also in this case, the LEI code may be missing or wrongly typed. In this respect, we remove all illicit LEI codes which are longer or shorter than 20 characters and those that include characters which are not letters or digits. This procedure is also applied to the SHS and AS security datasets. Nevertheless, these latter two datasets also have an ISIN dimension, which links multiple ISIN codes to the LEI code of the issuer. For the ISIN codes, as is the case for the LEI codes, they may be wrongly typed. In this respect, we remove all ISIN codes which are longer or shorter than their permitted length of 12 characters and those that include characters which are not letters or numbers. For those data entries in which the ISIN code is missing, we remove those entries from the datasets since it won't be possible to fill the missing ISIN during the mapping procedure even if we might have the related LEI code of the counterparty because multiple matches exist between ISIN codes to a single LEI. By contrast, we do keep those data entries with a missing LEI since only one unique match exists between an ISIN code and LEI code and between a NAME and LEI code.

1.2.2 Mapping Procedure

The key challenge of working with granular data is the unique identification of entities across reporting firms, datasets and time. The main objective of the mapping algorithm is to assign to each entity an identifier (ID number) which is unique across time. In order to achieve this outcome, we exploit the information reported by the firms regarding their exposures namely, the original counterparty NAME, the CLEAN_NAME we created, the LEI code and an ISIN code when it is a security exposure dataset (SHS and AS datasets). Moreover, in order to increase further the coverage of NAME, LEI codes and ISIN codes beyond those values already present in the exposure-datasets, we also retrieve two additional datasets from an open source website called GLEI[52]. The GLEI dataset has several dimensions and data tables. The GLEI_1 dataset provides a map of LEI to NAME of roughly 1.9 million entities[53]. The GLEI_2 dataset provides a map of ISIN to LEI of roughly 5.9 million securities[54]. These datasets are the most comprehensive map of legal classified entities around the world, and they are updated daily.

As first step of the procedure, we create a data table by combining all unique entities identified from each exposure-dataset. A unique entity is defined as one entity that has at least one field among the fields NAME, CLEANED NAME, LEI and ISIN which differs from another entity. If one field is missing (such as the ISIN field for non-security exposures), we exploit information only from the remaining fields. The same applies when one field is not available (NA), for instance because a reporting firm did not provide information about the LEI code of the counterparty entity. In this respect, we exploit information only by using the NAME and the CLEANED NAME variables. This initial data table is composed by roughly 8.6 million unique entities.

As a second step, we apply the mapping algorithm to the initial data table we created. The algorithm consists of a searching procedure which identifies sequentially all entities which display the same exact CLEAN_NAME across all entries, then the same LEI code and finally the same ISIN code and assigns to them a unique ID number. Hence, the algorithm looks for matches between pairs of fields: CLEAN_NAME to NAME, LEI to NAME, and ISIN to LEI. The rationale for the searching algorithm is as follows. One reporting firm may report the NAME of a given counterparty but not the related LEI code, while at the same time another reporting entity may report the NAME of the same counterparty along with the LEI code. The same logic applies to the other fields. Given this intuition, we are able to assign the same ID to the same entity by exploiting common information across pairs and across reporting firms, datasets and time periods. For the CLEAN_NAME to NAME matching pair, we use an exact string match rather than a fuzzy match since we want to avoid the risk of matching the wrong entity. We have to avoid introducing any matching error during the searching procedure because the error will propagate across matching pairs throughout the mapping procedure.

[52] Online Source: https://www.gleif.org/it/

[53] The data source is LEI-CDF v2.1 Golden Copy retrieved from: https://www.gleif.org/it/lei-data/gleif-golden-copy/download-the-golden-copy#/. Moreover, the GLEI golden source dataset also provides information on firms' attributes, such as geolocation, headquarter etc.

[54] The ISIN-LEI data source was retrieved from: https://www.gleif.org/it/lei-data/lei-mapping/download-isin-to-lei-relationship-files#

As a third step, once we have identified each entity with a unique ID, we create our final entity table by selecting one exemplar entity for each unique ID based on a statistical approach. Specifically, fields such as NAME and LEI codes for each unique ID will be then filled respectively with the NAME and LEI code that appear the most times for the same ID, that is, the most common entry NAME and LEI code. We do this, since sometimes firms report the wrong LEI code for the same NAME field. For this reason, during the searching procedure we give priority in assigning ID by NAME over LEI codes. In fact the searching procedure starts by assigning ID by NAME_CLEAN to NAME and after by LEI to NAME. The resulting Master Entity Table (Table 1) displays 2 million entries, roughly one fourth of the size of the initial data table.

Next, we add three additional columns to the table, respectively the COUNTRY, SECTOR and NACE fields of the entity, again by keeping the most common entry that appears in the data table for each variable. These dimensions provide relevant complementary information about the counterparty entity, which are important for modelling purposes—for instance to compute geographical or sectoral risk metrics (such as for stress testing exercises and for benchmarking aggregate exposures by country and sector).

Finally, we add to the entity table two additional columns referring to the consolidation status of the counterparty entity. This dimension is relevant since it gives us the flexibility to choose to model risk (depending on the scope of the exercise) at the highest level of consolidation or on an unconsolidated basis. By exploiting the GLEI_3 dataset, which provides a LEI relationship between each LEI code of the GLEI_1 dataset and its direct and ultimate parent company's LEI code, we construct two consolidation variables, respectively the ultimate parent ID within the UK jurisdiction (UP_ID_UK) and the ultimate parent ID worldwide (UP_ID_WW)[55]. To achieve that, we search across all LEI codes of the entity table, and once we find a match between the LEI code in the entity table and the GLEI_3 dataset, we assign the ID previously created matching the ultimate parent's LEI code reported in the GLEI_3 dataset. If the ultimate parent's LEI code in the GLEI-3 dataset matches the LEI code in the entity table, that entity is the ultimate parent and therefore the ID and the ultimate parent ID are the same. The distinction between ultimate parent ID within the UK and the ultimate parent ID worldwide is achieved by applying the same procedure but differentiating by the COUNTRY dimension of the entity. The mapping algorithm is computationally efficient, and

[55] The data sources is RR-CDF v.1.1 Golden Copy that can be retrieved from: https://www.gleif.org/it/lei-data/gleif-golden-copy/download-the-golden-copy#/

it takes roughly ~20 minutes to run. Moreover, the algorithm is scalable, that is, the addition of more information to the initial data table always improves its accuracy. Therefore by adding new granular data sources with the same dimensions and by increasing the time coverage of the datasets, the accuracy of the mapping function increases as long as the computational time. This is an important property of the function since the resulting data infrastructure can be automatically updated on a quarterly basis, and as a result, the mapping accuracy will improve endogenously and over time. To further improve the mapping accuracy, an ex-post manual cleaning exercise can be applied. The strategy is to select the top 100 exposures (in terms of amounts) not mapped after the first run of the algorithm and to construct a dataset with the same pair of fields and then manually fill in the LEI codes retrieved from the GLEI dataset. This dataset will be then introduced as an additional mapping source on top of the GLEI datasets into the mapping process. Thus, the information manually imputed will be automatically retrieved in following runs of the algorithm. The accuracy will improve non-linearly since the searching and mapping procedure will amplify the matching process via matching pairs. In this respect, Table 1 reports an example of the final output of the Master Entity Table. In particular, this example neatly shows how, through our mapping algorithm, we are able to successfully match an entity with its ultimate parent. We should emphasize that by having the LEI code dimension for each counterparty we can potentially add qualitative and quantitative variables to each counterparty by making matches with and exploiting other complementary data sources. Section 1.5 will provide additional information on the collection of complementary information for analytical and modelling purposes.

Table 1: Master Entity Table

NAME	LEI	ID	LEI UP	ID_UP	CTY	SEC	NACE
BP international limited	G1KG0OQD10NOMCMLDZ35	2469	213800LH1BZH3DI6G760	557	UK	NFC	G
BP plc	213800LH1BZH3DI6G760	557	/	557	UK	NFC	C

1.3 Efficiency of the Mapping Procedure

Using the Master Entity Data Table, we are able to assign our unique ID identifier to each counterparty in the original data source and to each reporting bank. Once we have implemented this step, we can refill the qualitative fields in the original datasets such as NAME, LEI, COUNTRY, SECTOR, and NACE classification with the cleaned information. This allows us to compute the efficiency of the mapping algorithm. In this respect, Table 2 provides a coverage of the LEI, COUNTRY, SECTOR and NACE dimensions before and after the mapping procedure. We want to emphasize that with our mapping approach we reach an LEI coverage above 87% across all granular data sources. This allows us to obtain a complete sectoral coverage and a very high country coverage above 93% of total exposure amounts. The sectoral classification by NACE economic activities is less accurate ranging between 62% and 92% of total coverage across data sources. This emphasizes the difficulties in aggregating information at a granular levels via counterparty-specific legal entity identifiers.

Table 2: Efficiency of the Mapping Algorithm by Data Source and Counterparty Dimension as share of total exposure coverage

Coverage				
LEI	LE	C67	SHS	AS
Raw Data	55%	95%	0%	66%
Clean Data	91%	97%	97%	87%
COUNTRY	LE	C67	SHS	AS
Raw Data	35%	91%	0	100%
Clean Data	93.00%	100%	98%	100%
SECTOR	LE	C67	SHS	AS
Raw Data	35%	91%	0	0
Clean Data	100%	100%	100%	100%
NACE	LE	C67	SHS	AS
Raw Data	22%	0	0	0
Clean Data	85%	90%	92%	62%

1.4 Data Infrastructure

We construct a data infrastructure in order to automatize the mapping procedure and make the dataset easily updatable so as to extend the horizon of the Global Network on a quarterly basis, and in doing so regularly improve its efficiency. The data infrastructure has been developed in R and consists of three main blocks. The first block concerns the automatic download and preparation of the original data sources in a standardized format, that is, by selecting only the dimensions/columns previously described. Since we are working with big granular datasets, reducing the dimension of the data frames is crucial to minimize the computational time of the whole process. Next, we apply the cleaning and mapping algorithm in order to consistently identify entities with unique IDs. This block is fundamental to be able to uniquely identify exposures across datasets between the same reporting firm and counterparty entities, and in doing so, also be able to subset the datasets consistently thereby avoiding double-counting of exposures across datasets during the merging procedure. Therefore, the higher the accuracy of the mapping algorithm, the higher is the accuracy of the merging procedure. Finally, the third block aims at constructing the Global Network which will be used for policy and research purposes and is divided into the three following steps.

1.4.1 Step 1 - Consolidate Reporting Side

The first step involves the selection of reporting firms. To avoid double counting of exposures between subsidiaries and the head of group, we select only those UK reporting banks that report at the highest consolidation level. In this respect, we select a subset of firms, respectively consolidated banking groups and solo entities which do not belong to any other consolidated banking group. In this subset, two types of firms will appear, UK banks and non-UK bank subsidiaries domiciled in the UK. For those datasets which only report exposures at a subsidiary level such as the AS security dataset, by exploiting the information collected on the GROUP structure, we will consolidate the reporting side by merging security exposures of multiple subsidiaries belonging to the same banking group into a single entity. This would allow us to display on the reporting side the same banking groups and solo entities across datasets. The objective of the manipulation is threefold: i) homogenize the reporting side; ii) further reduce the size of the datasets; iii) avoid double counting of reported exposures.

1.4.2 Step 2 - Exposure Cleaning and Refilling

Since firms sometimes make mistakes on the original reported amount of the exposures, we run some quality control checks. We remove all exposures that, as gross original amounts, are larger than 20% of a firm's total assets from the LE and C67 datasets and those security exposures (ISIN based) that are larger than 6% of a firm's total assets. The selection of these thresholds is an ad-hoc selection based on our specific-reported sample. Second, to reduce further the size of the datasets we set a floor cap to the size of the exposure. We keep all exposures that are larger than £1 million for the LE and the C67 datasets. As far as the security datasets are concerned, since they also report *short* exposures which are reported as negative values, we keep all exposures that are larger than £1 million in absolute value, i.e. > £1 million and < –£1 million. In the end, once we have prepared the reporting side and cleaned exposure amounts that are considered as outliers, we proceed to refill the security datasets from missing quarters so as to achieve a homogeneous quarterly frequency across data sources.

The LE dataset is reported with a quarterly frequency, and the sample starts in Q1-2015. The C67 dataset is reported with a monthly frequency, and to align the two datasets, we select exposures reported on the months that match the end of each reporting quarter. These two datasets are our core data sources and they do not require any further data manipulation in this respect. By contrast, the SHS dataset has an annual frequency with a snapshot date based in Q4, starting from 2018. Whereas, the AS dataset has a quarterly frequency, but the collection started only in Q1-2019. Hence, we exploit an interesting characteristic of these granular exposure datasets. Overall, we notice that exposures among a reporting firm and a counterparty entity tend to be constant over time, with only small variations in terms of exposure amounts across quarters. For instance exposure amounts at T-2 and T+2 are relatively similar. This shows that the largest exposures in our datasets are slow-moving, and the larger is the exposure amount, the higher is the likelihood of this feature being valid. This characteristic is evident in both exposure-based and security-based datasets. We exploit this property of our data and we fill the missing quarters of the SHS datasets (Q1-Q2-Q3) by back-filling the values provided for the Q4 snapshot. For those security contracts which are debt contracts, we also have to adjust the maturity and issuance date, which is recalculated proportionally to the time shift. The same filling strategy is applied to the AS dataset thereby bringing backwards the Q1-2019 snapshot to fill the missing snapshots in Q4, Q3, Q2, and Q1 of 2018. We do this mainly to extend the time coverage of the datasets in the past to the largest extent possible.

Finally, before merging exposures consistently across datasets and creating the Global Network of Granular Exposures, we implement a final step in which we create a common data-column structure across datasets and fill in values for entries in rows and columns which present NAs. Every dataset should have nine columns set out as follows. Column (1) for the reporting bank ID (also defined as LENDER), column (2) for the counterparty entity ID (also defined as BORROWER), column (3) identifying the REPORTED_PERIOD, column (4) identifying the type of exposures called SEC_TAG, column (5) reporting the original exposure amount (also defined ORIGINAL_EXP), column (6) reporting the short-term exposure amount (also defined ORIGINAL_EXP_ST), column (7) reporting the equity exposure amount (also defined EQUITY), column (8) reporting the net exposure amount (also defined NET_EXP), and column (9) reporting the source of the dataset (also defined SOURCE)[56].

The LE dataset represents our core dataset and the column-structure described above resembles the one already present in the LE dataset. The columns that show exposure amounts are columns (5) to (9). Column (5), which shows gross exposure amounts, is completely filled with values in every dataset. However, we have missing values for column (6), that is, short-term exposures for the LE dataset, since the maturity structure of exposures is provided only

[56] Some datasets have some empty columns which are fill with 0. For instance, the C67 dataset has no equity exposures, thereby column (7) will be filled with zeros. Contrary the security datasets do not have non-security exposures by constructions, and also it will be filled with zeros.

for the top-20 exposures in the LE dataset. Hence, we fill short-term exposure amounts for these rows using the average short-term exposure amount calculated over time for each lender against all relevant borrowers. If NAs are still present, we use the average short-term exposure amount by counterparty borrower, or otherwise the average by country and counterparty sector. Using this approach, we are able to fill all short-term exposure amounts for rows corresponding to the LE dataset. Another column that exhibits NAs is column (8) "net exposure amounts" capturing unsecured exposures in the C67 dataset, as that column is missing from the original data source. In this case, we exploit information on the type of instrument, which is provided in the C67 dataset. Hence, we insert 0 when an instrument is a collateralized instrument[57]. Otherwise, we insert the average percentage by reporting and counterparty entity calculated from the LE dataset. The same procedure is applied to the AS and SHS security datasets for which column (8) is also missing. Finally, column (7) is missing from the C67 dataset. However, in this case we directly insert 0 for these rows against this column since in the C67 datasets there are no equity instruments.

1.4.3 Step 3 - Merging Strategy

The final step consists in two consolidation procedures: i) aggregation of exposures by counterparty entity; and ii) merging of exposures consistently across datasets. The former aims at aggregating exposures of one reporting firm within the same time period vis-à-vis multiple counterparty entities belonging to the same group of entities, whether they are credit institutions, financial corporations or non-financial corporations. Since we have reconstructed the group structure for each counterparty entity, we can use the variable UK_UP_ID to sum exposure amounts across counterparty entities belonging to the same ultimate parent within the UK jurisdiction for rows corresponding to the LE and C67 datasets. The fields we sum are the gross exposure amounts, the short-term exposure amounts, the net exposure amounts, and the equity exposure amounts. Similarly, we implement the same procedure for the security datasets by summing all amounts by ISIN codes for the same counterparty entity although differentiating between equity and debt security exposures. We then aggregate exposures by UK_UP_ID of the counterparty as we had done in the case of the LE and C67 datasets. Therefore we may have two security exposures between a reporting firm and the same counterparty entity within the same time period, respectively one for equity and one for debt instruments. Hence, the security datasets will lose the ISIN dimension, and they become

[57] The set of instruments is the following: funding obtained from intragroup counterparties (IGCP), funding obtained from repurchase agreements (SFT), other secured wholesale funding (OSWF), and other funding products (OFP). Info about the instruments can be found here: eba.europa

exposure-based datasets, that is, displaying an entity to entity relationship identified by a LEI code. Nevertheless, the ISIN-related information won't be lost since we keep a column for short-term exposure amounts, which are calculated by summing exposure amounts across debt security contracts with maturity date below 30 days. Moreover we keep a clear distinction between equity and debt exposures. Once we have consolidated the exposures and security exposures by UK_UP_ID, we have four datasets with the very same column and row structure.

The second step of the procedure consists of merging exposures across datasets. We first construct the network of LE and C67 exposures, also defined as edges_LE_C67. To do this, we add to the LE dataset all exposures from the C67 dataset that are not already present in the LE dataset. Secondly, when an exposure between the same reporting firm and counterparty entity appears in both LE and C67, we keep the exposure reported in the LE dataset. Next, we implement the same procedure with the security datasets SHS and AS. We derive the network of securities (edges_AS_SHS) by adding to the AS dataset all exposures from the SHS dataset that are not already present in in the AS dataset. Hence, we give priority to the AS dataset over the SHS dataset since the AS dataset has a quarterly frequency, as we have already noted. Finally, we merge the two newly created networks, that is, edges_LE_C67 and the edges_AS_SHS. Before doing so, we split the edges_LE_C67 exposures into two sets. We subtract the equity exposures from the gross original exposure amounts and we add the equity exposures as new rows into the dataset and we distinguish these rows in the SOURCE column by tagging them as "SLE". These exposures are equity security exposures stemming from the LE dataset. Hence, the edges_LE_C67 network is divided into debt exposures (loan + security) and equity security exposures. Next, we aim to separate loan exposures in the edges_LE_C67 network from those exposures that instead are debt security exposures. To achieve that, we match debt security exposures among the same reporting firm and counterparty entity among the two network datasets edges_LE_C67 and edges_AS_SHS for each reported period. Then we remove from the edges_LE_C67 network the gross exposure amount that appears in the edges_AS_SHS security network. Hence, the remaining amount of debt exposures tagged as "LE" and "C67" in the SOURCE column are now only loan exposures. In this respect, in the SEC_TAG column we now classified those debt exposures as loan exposures "L". Then we add those exposures from the edges_AS_SHS network to the edges_LE_C67 dataset as new rows. As was the case for exposure values derived from the LE dataset, the SOURCE column will also identify the original data sources from which those rows come from, such as AS or SHS datasets. We then carry out the same procedure for those matched equity security exposures that match exposures in the LE dataset and we keep those that are stemming from the LE dataset[58]. If the matched equity exposure from the edges_AS_SHS dataset is larger than the exposure from the edges_LE_C67 dataset, we add an additional row with an amount that is equal to the difference in the two amounts. In contrast, all exposures that appear in the edges_AS_SHS network and do not appear in the edges_LE_C67 network (unique exposures) are directly added as new rows. We thus arrive at the completion of the Global Network of Granular Exposures which is made of i) loan exposures, ii) debt security exposures, and iii) equity security exposures. Nevertheless, the sum across all exposures for each reporting bank may not still match the total assets of each bank for two main reasons: i) loan exposures to the household sector, and ii) loan exposures towards small-medium size enterprises which are not captured by the LE dataset since they are smaller than LE reporting threshold. In this respect, in order to complete the asset side coverage we exploit the supervisory template F.20.04 reporting aggregate exposures by country and sector of counterparty. First we assign a unique ID to each country-sector pair starting from the last sequential ID number previously defined in the entity table and we add them to the entity table as new counterparty entities. The name assigned to these rows will be defined as the combination of the country and sectoral codes, for instance UK-HH for the UK household sector. Then, we aggregate granular exposures from the Global Network by reporting firm and reported period and by sector and country pair. We then match for each reporting firm the total aggregate exposures by country and sector retrieved from the supervisory template F.20.04 with the aggregated granular exposures by country and sector computed from the Global Network. Hence, for each reporting firm and reported period and country and sector pair we take the difference in the two exposure amounts (aggregate exposures – aggregated granular exposures) and we add it as new rows. These new rows are then identified under the SOURCE column with the tag "FIN" from FINREP data source and we classify them as loan exposures "L" in the column SEC_TAG. The Global Network of Granular Exposures is thus augmented with aggregate exposures to each country-sector pair. Table 3 reports the number of reporting banks we cover for each single data source after these procedures have been applied. We have in total 36 banks, out of which 18 banks have a complete coverage across LE, SEC and FINREP datasets.

Table 3: Coverage of Reporting Banks by Dataset

Reporting Banks	**LE**	**SEC**	**FINREP**	**COMPLETE**
36	21	25	18	18

[58] The amount of equity exposures here matched is a small share of the total equity exposure amounts.

1.5 Complementary Data Sources

We now provide additional information on other complementary data sources we use for modelling and stress testing purposes. First of all, we collect solvency information on banks' capital base (CET1), risk weighted assets (RWAs), and total assets (TA), respectively from COREP supervisory templates C.01 and C.02 as well as FINREP supervisory template F.01. Then we complement exposure and counterparty information by retrieving loss given default (LGD) and probability of default (PD) parameters from COREP supervisory templates C.09.02 and approximating these values to apply by country and sector of the counterparty[59]. This template is submitted by firms who have received prior approval from the PRA to use the internal-ratings based (IRB) approach to determine capital requirements for certain exposures and whose non-domestic exposures are greater than 10% of total exposures. Under the IRB approach, firms are allowed to rely on their own estimates of risk components such as PD and/or LGD rather than using supervisory estimates for most asset classes in scope. In general, in order to arrive at these estimates, banks need to build a model that meaningfully differentiates risk by defining grades and assigning obligors or exposures to each grade or pool.

The COREP supervisory templates C.09.02 provides a detailed breakdown of each reporting bank' LGDs and PDs parameters by country and sector of the counterparty. Firms provide these estimates for 5 major obligor sectors such as corporate, sovereign, bank, retail and equity[60]. PD parameters are calibrated to the long-run average PD of one-year default rates based on exposures on the banking book. While PDs are based at the obligor level, LGDs are based at the facility level. Overall, the PD and LGD datasets covers on a quarterly basis roughly 17 reporting banks at the highest level of consolidation, 247 countries, and obligors in 16 sub-sectors, for a total of 485.834 data points from Q1-2018 up to Q4-2021. In this respect, as an illustrative example, Panel A and Panel B of Table 4 reports the development over time of the average PD and LGD parameters across sectors and by regional location of the counterparty. These estimates are the result of a pool-estimation approach since we compute an exposure-weighted average across all reporting banks in the sample. We want to stress the fact that these parameters are an average across countries, and strong heterogeneity exists among countries belonging to the same region. We also note that while these estimates are based on supervisory templates submitted by a limited sub-set of reporting banks, this dataset provides useful

[59] The structure of the COREP templates can be retrieved here: eba.europa

[60] The corporate asset class can be also split between exposures towards SMEs and non-SMEs and other sub-classes. A similar splitting is applied to retail exposures. For an overview of the sector and asset classes covered see: CRE30 - IRB approach: overview and asset class definitions (bis.org)

insights into firms' internal estimates of risk by country and sector albeit at an aggregate level. In theory, the optimal approach would be to map each entity with one ad-hoc counterparty-specific PD parameter. However, datasets like Moodys Riskcalc, though they provide quarterly time series of PD and LGD parameters by firm identified with an LEI code, these datasets cover only a sub-set of large corporates[61].

Table 4 – Panel A: Average PDs by Sector and Region

Region-Sector	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
Africa_CB	3.0%	3.5%	2.9%	3.0%	3.2%	3.2%	3.3%	3.6%	2.7%	2.9%	2.8%	3.1%	3.7%	3.5%	3.3%	3.5%
Americas_CB	0.3%	0.3%	0.4%	0.4%	0.4%	1.1%	0.9%	0.9%	1.0%	0.3%	0.4%	0.4%	0.5%	0.5%	1.1%	1.1%
Asia_CB	0.5%	0.5%	0.6%	0.6%	0.6%	0.6%	0.6%	0.5%	0.8%	0.6%	0.7%	0.9%	0.9%	0.9%	1.0%	0.9%
Europe_CB	0.1%	0.4%	0.2%	0.3%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.3%	0.3%
Oceania_CB	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
UK_CB	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	0.1%
Africa_CI	2.5%	1.9%	2.2%	1.9%	2.4%	2.6%	1.7%	2.4%	2.1%	1.5%	1.7%	2.4%	3.0%	3.9%	3.7%	2.5%
Americas_CI	2.1%	2.0%	1.0%	1.2%	1.3%	1.0%	1.3%	1.5%	0.8%	0.6%	0.7%	0.9%	0.7%	0.7%	1.2%	1.3%
Asia_CI	1.0%	0.9%	1.0%	1.0%	1.0%	1.1%	1.1%	1.2%	1.0%	1.0%	1.3%	1.5%	1.6%	1.6%	1.8%	1.8%
Europe_CI	0.6%	0.7%	0.7%	0.7%	0.7%	0.7%	0.8%	0.7%	0.7%	0.7%	0.8%	0.8%	0.8%	0.5%	0.7%	0.6%
Oceania_CI	0.1%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	0.1%
UK_CI	0.4%	0.4%	0.3%	0.3%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.2%	0.1%	0.1%	0.1%	0.1%
Africa_FC	4.3%	5.4%	3.7%	4.7%	4.2%	4.9%	5.8%	4.9%	4.8%	4.9%	5.7%	5.9%	6.4%	6.2%	6.9%	5.4%
Americas_FC	1.8%	2.0%	1.9%	2.4%	2.4%	2.4%	2.1%	2.0%	1.9%	1.8%	2.0%	1.8%	1.7%	1.6%	1.6%	1.6%
Asia_FC	2.5%	2.4%	2.6%	2.6%	2.5%	2.5%	2.5%	2.6%	2.3%	2.9%	2.8%	3.2%	3.1%	3.0%	3.1%	2.9%
Europe_FC	1.1%	1.1%	1.0%	1.6%	1.0%	1.2%	1.1%	1.4%	1.6%	1.6%	1.6%	1.5%	1.5%	1.1%	1.1%	1.0%
Oceania_FC	0.6%	0.8%	0.7%	0.7%	0.7%	1.0%	1.8%	0.9%	1.0%	1.2%	1.3%	1.5%	1.1%	1.2%	0.9%	0.9%
UK_FC	2.3%	2.3%	2.1%	2.1%	2.2%	2.3%	2.2%	2.2%	2.3%	2.7%	2.8%	2.9%	2.7%	2.6%	2.5%	2.2%
Africa_GG	3.0%	3.5%	2.9%	3.0%	3.2%	3.2%	3.3%	3.6%	2.7%	2.9%	2.8%	3.1%	3.7%	3.5%	3.3%	3.5%
Americas_GG	0.3%	0.3%	0.4%	0.4%	0.4%	1.1%	0.9%	0.9%	1.0%	0.3%	0.4%	0.4%	0.5%	0.5%	1.1%	1.1%
Asia_GG	0.5%	0.5%	0.6%	0.6%	0.6%	0.6%	0.6%	0.5%	0.8%	0.6%	0.7%	0.9%	0.9%	0.9%	1.0%	0.9%
Europe_GG	0.1%	0.4%	0.2%	0.3%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.3%	0.3%
Oceania_GG	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
UK_GG	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	0.1%
Africa_HH	0.6%	0.6%	0.6%	0.5%	0.5%	0.6%	0.6%	0.5%	0.6%	0.6%	0.6%	0.6%	0.7%	0.6%	0.5%	0.5%
Americas_HH	0.3%	0.6%	0.6%	0.6%	0.6%	0.5%	0.5%	0.4%	0.5%	0.5%	0.6%	0.6%	0.6%	0.6%	0.6%	0.4%
Asia_HH	0.5%	0.8%	0.7%	0.6%	0.6%	0.6%	0.6%	0.7%	0.7%	0.7%	0.9%	1.0%	0.9%	0.8%	0.9%	0.9%
Europe_HH	0.7%	0.9%	0.9%	0.8%	0.8%	0.7%	0.9%	0.9%	0.9%	1.0%	0.8%	1.0%	0.9%	0.8%	0.9%	0.7%
Oceania_HH	0.4%	0.5%	0.5%	0.2%	0.3%	0.2%	0.2%	0.6%	0.6%	0.5%	0.6%	0.5%	0.7%	0.4%	0.2%	0.3%
UK_HH	1.1%	1.1%	1.2%	1.3%	1.2%	1.3%	1.3%	1.3%	1.2%	1.4%	1.3%	1.4%	1.4%	1.2%	1.2%	1.1%
Africa_NFC	5.1%	6.7%	4.5%	5.9%	4.9%	5.4%	6.3%	5.4%	5.3%	5.9%	6.6%	6.9%	7.2%	6.7%	7.3%	6.1%
Americas_NFC	1.8%	1.6%	1.9%	2.3%	2.2%	2.4%	2.2%	2.1%	2.1%	1.9%	2.2%	2.1%	1.8%	1.8%	1.9%	1.9%
Asia_NFC	2.9%	2.8%	3.1%	3.0%	3.0%	2.9%	2.9%	3.0%	2.6%	3.6%	3.6%	3.9%	3.5%	3.4%	3.4%	3.3%
Europe_NFC	1.5%	1.5%	1.3%	2.0%	1.2%	1.6%	1.2%	1.5%	2.0%	2.0%	2.0%	1.9%	1.9%	1.6%	1.5%	1.6%
Oceania_NFC	0.8%	1.0%	0.9%	0.8%	0.8%	1.1%	1.9%	1.0%	1.2%	1.3%	1.4%	1.8%	1.2%	1.2%	1.0%	1.0%
UK_NFC	2.7%	2.6%	2.4%	2.4%	2.6%	2.7%	2.6%	2.6%	2.6%	3.1%	3.2%	3.4%	3.2%	3.1%	3.0%	2.7%

Table 4 – Panel B: Average LGDs by Sector and Region

[61] Information on Riskcalc can be retrieved at the following website.

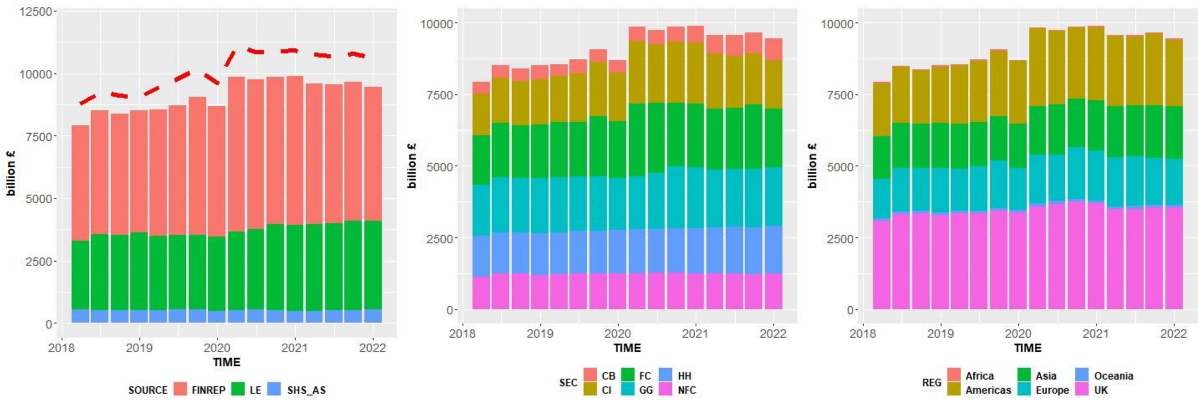
Region-Sector	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
Africa_CB	45%	46%	46%	45%	45%	44%	44%	43%	44%	44%	44%	44%	45%	44%	45%	44%
Americas_CB	40%	40%	41%	41%	41%	42%	42%	42%	42%	42%	42%	40%	40%	44%	44%	45%
Asia_CB	45%	46%	45%	45%	44%	45%	44%	45%	45%	45%	45%	45%	45%	44%	44%	44%
Europe_CB	54%	54%	47%	47%	47%	47%	46%	47%	47%	46%	46%	46%	46%	46%	46%	47%
Oceania_CB	41%	40%	39%	44%	41%	40%	40%	44%	43%	42%	44%	46%	46%	45%	45%	45%
UK_CB	56%	61%	48%	49%	47%	47%	46%	48%	48%	49%	48%	49%	49%	48%	48%	48%
Africa_CI	42%	43%	41%	40%	38%	38%	37%	38%	36%	32%	35%	28%	34%	37%	36%	33%
Americas_CI	41%	43%	41%	42%	40%	41%	43%	43%	42%	39%	38%	36%	36%	37%	35%	33%
Asia_CI	39%	38%	36%	36%	36%	36%	36%	37%	37%	36%	34%	35%	36%	35%	33%	33%
Europe_CI	53%	53%	49%	51%	49%	51%	49%	49%	50%	49%	47%	49%	49%	50%	46%	46%
Oceania_CI	45%	43%	49%	45%	47%	46%	45%	46%	45%	44%	43%	44%	42%	43%	42%	40%
UK_CI	41%	43%	39%	37%	38%	37%	37%	35%	37%	38%	36%	37%	35%	34%	32%	35%
Africa_FC	42%	41%	44%	43%	44%	44%	43%	45%	42%	42%	41%	39%	41%	42%	41%	40%
Americas_FC	36%	40%	41%	43%	41%	42%	43%	40%	41%	38%	37%	36%	37%	35%	38%	37%
Asia_FC	41%	41%	41%	40%	41%	41%	42%	43%	42%	41%	40%	41%	41%	40%	38%	39%
Europe_FC	48%	49%	46%	48%	47%	49%	48%	47%	49%	47%	44%	43%	43%	45%	45%	44%
Oceania_FC	46%	45%	47%	45%	47%	46%	48%	48%	46%	46%	46%	45%	40%	40%	38%	39%
UK_FC	39%	40%	38%	37%	38%	38%	37%	37%	38%	37%	36%	36%	35%	37%	35%	36%
Africa_GG	45%	46%	46%	45%	45%	44%	44%	43%	44%	44%	44%	44%	45%	44%	45%	44%
Americas_GG	40%	40%	41%	41%	41%	42%	42%	42%	42%	42%	42%	40%	40%	44%	44%	45%
Asia_GG	45%	46%	45%	45%	44%	45%	44%	45%	45%	45%	45%	45%	45%	44%	44%	44%
Europe_GG	54%	54%	47%	47%	47%	47%	46%	47%	47%	46%	46%	46%	46%	46%	46%	47%
Oceania_GG	41%	40%	39%	44%	41%	40%	40%	44%	43%	42%	44%	46%	46%	45%	45%	45%
UK_GG	56%	61%	48%	49%	47%	47%	46%	48%	48%	49%	48%	49%	49%	48%	48%	48%
Africa_HH	38%	43%	44%	43%	43%	43%	43%	44%	40%	40%	41%	39%	41%	41%	40%	40%
Americas_HH	33%	34%	35%	36%	36%	38%	38%	36%	39%	36%	35%	34%	34%	33%	36%	35%
Asia_HH	42%	39%	39%	39%	40%	40%	40%	41%	41%	38%	37%	38%	37%	37%	36%	36%
Europe_HH	48%	48%	44%	44%	45%	46%	45%	45%	46%	45%	42%	42%	41%	43%	43%	43%
Oceania_HH	39%	34%	41%	33%	35%	38%	39%	38%	34%	37%	37%	39%	38%	38%	39%	39%
UK_HH	29%	29%	29%	28%	29%	29%	28%	28%	28%	28%	27%	27%	26%	27%	26%	26%
Africa_NFC	41%	39%	43%	42%	43%	44%	43%	45%	42%	44%	42%	41%	42%	42%	42%	41%
Americas_NFC	35%	38%	42%	42%	41%	41%	42%	40%	42%	41%	39%	40%	41%	39%	40%	40%
Asia_NFC	43%	42%	42%	42%	43%	42%	45%	45%	44%	43%	43%	43%	43%	43%	42%	42%
Europe_NFC	46%	47%	45%	47%	46%	47%	47%	46%	48%	47%	44%	41%	41%	43%	44%	42%
Oceania_NFC	46%	45%	45%	45%	47%	46%	48%	48%	46%	46%	46%	45%	40%	40%	38%	39%
UK_NFC	39%	39%	38%	38%	38%	38%	37%	37%	38%	37%	36%	36%	36%	37%	36%	36%

1.6 Multilayer Network Statistics

The Global Network of UK banks' exposures is composed by six data sources divided into two loan exposure datasets (LE, FINREP) and three security exposures datasets (SHS, SAS, SLE). There is one variable of interest, namely, gross original exposure amounts, although we also collect net exposure amounts which can be a proxy for exposures at default (EAD).

Figure 1 presents the coverage over time of the Global Network by data source and by type of exposure. Panel (a) compares the total amount of UK banks' gross exposures with the UK banking sector's total assets (red dotted line) at each point in time. The Global Network captures £9.4 trillion of exposures out of £10.6 trillion of total assets in Q4-2021, roughly 90% of the UK banking system' asset side. The exposure coverage relative to total assets is stable over time. The average quarter-on-quarter variation is equal to 3 percentage points, with a maximum variation of 9% in Q1-2020 relative to Q4-2019 consistent with a similar expansion of the UK banking sector's balance sheet. This feature also holds for the contribution of each dataset to the total coverage, which is stable over time. The aggregated exposures dataset by country and sector (derived from FINREP) contributes to 57% of the total coverage, whereas all granular exposure datasets combined make 43%. On the one hand, security exposure datasets (SHS_AS) capture 6% of the total[62]. On the other hand, granular loan exposure datasets capture 33% of the total, while granular derivative exposures captures roughly 5%. Figure 1 thus highlights the rationale and value-add of combining and exploiting the various data sources described in section 1. The outcome is a good and stable coverage of the UK banking sector's asset side. Having detailed the contribution of each data sources to the composition of the Global Network, we now move to the presentation of the decomposition by counterparty sector and country of origin of UK banks' exposures. We show that in Q4-2021, the most relevant counterparty sector is governments (GG) capturing 21.8% of total gross exposure amounts. Then follow exposures to financial corporations (FC) with 21.6%, and, after that, exposures to credit institutions (CI) with 18%, to the household sectors (HH) with 17.4%, to non-financial corporates (NFC) with 13.3%, and finally to central banks (CB) with 8%. In the end, the contribution over time by geographical region (We do not provide the breakdown by country for graphical purposes.) highlights that in Q4-2021, UK banks were mostly exposed outside the UK for 62.3% of total gross exposures, while domestically for the remaining 37.6%. The most relevant regions outside the UK are Americas with 24.5%, Asian countries with 19.5%, and Europe with 16.9%.

Figure 1: Time-Series Decomposition of total gross exposure coverage by data source, sector and country



Note: red dotted line refer to UK banks' total assets.

Appendix B – Conditional Capital at Risk Estimates

Table 1 – Decomposition of UK Banks' CCaR Estimates (Extreme Stress Scenarios)

Panel (a) – By Sector

[62] Sydow et al. (2021) provides estimate for the share of EA banks' granular loans to individual firms, which represents roughly 21% of banks' total assets. Moreover, they highlight that information on granular securities holdings in banks' balance sheets covers only 7% of total assets.

Loss Ratio	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
CI	0.3%	0.3%	0.2%	0.2%	0.3%	0.3%	0.3%	0.2%	0.2%	0.3%	0.3%	0.2%	0.2%	0.3%	0.2%	0.2%
GG	1.0%	1.1%	1.5%	1.8%	1.5%	1.1%	1.0%	1.0%	0.8%	1.0%	1.1%	0.9%	0.9%	0.7%	1.2%	1.1%
FC	11.3%	11.2%	10.0%	10.2%	10.3%	10.1%	10.5%	12.5%	12.8%	15.5%	11.7%	12.2%	10.9%	10.9%	10.2%	10.7%
NFC	7.1%	7.6%	6.9%	7.2%	7.4%	7.8%	8.3%	6.8%	7.3%	8.3%	8.9%	9.5%	9.0%	9.1%	8.5%	8.6%

Loss Share	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
CI	1.3%	1.4%	1.0%	1.0%	1.1%	1.2%	1.2%	0.7%	1.0%	0.8%	1.3%	1.1%	1.2%	1.4%	1.1%	1.0%
GG	14.0%	14.7%	21.5%	24.4%	19.7%	15.3%	13.6%	11.8%	8.9%	9.4%	14.5%	12.9%	13.2%	10.4%	17.2%	15.8%
FC	53.8%	51.5%	47.5%	45.0%	47.9%	48.6%	47.8%	62.2%	64.6%	65.9%	52.2%	52.2%	49.9%	52.0%	49.7%	50.6%
NFC	30.9%	32.4%	30.0%	29.6%	31.2%	35.0%	37.4%	25.3%	25.4%	23.9%	32.0%	33.9%	35.7%	36.2%	32.0%	32.6%

Panel (b) – By Region

Loss Ratio	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
UK	6.7%	6.4%	6.3%	6.7%	6.1%	5.7%	5.7%	8.1%	8.3%	9.2%	5.4%	5.4%	5.5%	5.6%	5.9%	5.4%
Americas	3.2%	3.2%	2.8%	3.1%	3.3%	3.0%	3.7%	2.7%	2.9%	3.6%	4.2%	4.2%	3.9%	3.8%	4.0%	4.0%
Europe	3.2%	3.4%	3.2%	3.2%	3.5%	3.8%	3.5%	2.8%	2.9%	3.7%	3.5%	3.6%	3.1%	3.1%	3.0%	3.1%
Asia	1.0%	0.9%	0.9%	1.0%	0.9%	0.9%	1.0%	0.8%	0.9%	1.1%	1.2%	1.1%	1.0%	1.1%	1.1%	1.3%
Africa	2.5%	4.4%	3.1%	3.7%	3.1%	3.2%	4.5%	3.3%	2.7%	4.5%	6.0%	5.8%	4.6%	5.0%	5.8%	6.6%
Oceania	0.8%	0.8%	0.7%	0.7%	0.9%	0.9%	1.0%	0.6%	0.7%	1.0%	1.1%	1.0%	0.8%	1.0%	0.7%	0.9%

Loss Share	Q1-2018	Q2-2018	Q3-2018	Q4-2018	Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020	Q2-2020	Q3-2020	Q4-2020	Q1-2021	Q2-2021	Q3-2021	Q4-2021
UK	55.4%	55.6%	58.0%	57.6%	55.7%	51.8%	50.3%	67.4%	64.5%	64.7%	48.0%	46.4%	46.4%	46.2%	48.1%	47.0%
Americas	15.8%	15.5%	14.1%	15.6%	16.9%	16.0%	19.4%	12.7%	13.7%	13.7%	20.7%	20.9%	22.1%	21.9%	23.4%	22.4%
Europe	22.1%	22.6%	21.8%	19.9%	21.9%	25.7%	23.2%	15.1%	16.2%	16.2%	23.8%	24.6%	23.2%	22.9%	19.7%	20.8%
Asia	6.0%	5.4%	5.4%	6.1%	4.9%	5.6%	6.0%	4.2%	5.0%	4.6%	6.4%	6.9%	7.3%	8.0%	7.8%	8.7%
Africa	0.4%	0.7%	0.5%	0.6%	0.5%	0.5%	0.8%	0.5%	0.4%	0.5%	0.9%	0.9%	0.8%	0.8%	0.9%	0.8%
Oceania	0.3%	0.2%	0.2%	0.2%	0.3%	0.3%	0.3%	0.1%	0.2%	0.2%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%

Note: The "Loss Ratio" is computed as losses over gross exposures by country or region, while the Loss Share is calculated as sectoral/regional losses over total losses.